

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OMprotein - protein search, using sw model

Run on: June 29, 2004, 10:22:54 ; Search time 18 Seconds
(without alignments)
2846.502 Million cell updates/sec

Title: US-10-024-368-2

Perfect score: 5358

Sequence: 1 FYYKDPSIPRLAKEPLAAAE.....MDWAVFQAVKVAVGTQLQEA 984

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5350	99.9	1189	1	HAIR_HUMAN	O43593 homo sapien
2	4110.5	76.7	1182	1	HAIR_MOUSE	Q61645 mus musculu
3	4067	75.9	1181	1	HAIR_RAT	P97609 rattus norv
4	694.5	13.0	1214	1	TSGA_RAT	Q63679 rattus norv
5	205	3.8	1616	1	APXL_HUMAN	Q13796 homo sapien
6	190.5	3.6	660	1	YHL1_EBV	P03181 epstein-bar
7	189	3.5	2167	1	SHK1_RAT	Q9wv48 rattus norv
8	186.5	3.5	1461	1	IE18_PRVIF	P11675 pseudorabie
9	183	3.4	1446	1	IE18_PRVKA	P33479 pseudorabie
10	182.5	3.4	1723	1	KA93_HUMAN	Q9upq9 homo sapien
11	180.5	3.4	2161	1	SHK1_HUMAN	Q9y566 homo sapien
12	180	3.4	1733	1	VNUA_PRVKA	P33485 pseudorabie
13	176	3.3	2774	1	MAPA_RAT	P34926 rattus norv
14	173	3.2	1783	1	RAA3_CHLRE	Q9fec4 chlamydomon
15	172	3.2	1670	1	CA34_HUMAN	Q01955 homo sapien
16	171.5	3.2	2142	1	BAT2_HUMAN	P48634 homo sapien
17	170.5	3.2	5262	1	MLL2_HUMAN	O14686 homo sapien
18	168.5	3.1	1466	1	CA13_HUMAN	P02461 homo sapien
19	167.5	3.1	1017	1	EM11_MOUSE	Q99k41 mus musculu
20	166.5	3.1	1143	1	CA1I_HUMAN	Q14993 homo sapien
21	166	3.1	1372	1	CA21_MOUSE	Q01149 mus musculu
22	164	3.1	1464	1	CA13_MOUSE	P08121 mus musculu
23	163.5	3.1	1411	1	TCOF_HUMAN	Q13428 homo sapien
24	163.5	3.1	1459	1	CA12_MOUSE	P28481 mus musculu
25	163	3.0	1366	1	CA21_HUMAN	P08123 homo sapien
26	162	3.0	1690	1	CA44_HUMAN	P53420 homo sapien
27	162	3.0	1691	1	CA64_HUMAN	Q14031 homo sapien
28	162	3.0	2944	1	CA17_HUMAN	Q02388 homo sapien
29	161.5	3.0	1603	1	CA1F_HUMAN	Q07092 homo sapien
30	160	3.0	1262	1	CA13_CHICK	P12105 gallus gall
31	160	3.0	1366	1	CA21_CANFA	O46392 canis famil
32	159	3.0	1372	1	CA21_RAT	P02466 rattus norv
33	159	3.0	1418	1	CA12_HUMAN	P02458 homo sapien

RESULT 1
HAIR_HUMAN
ID HAIR_HUMAN STANDARD; PRT; 1189 AA.
AC O43593; Q96H33; Q9NPE1;
DT 15-JUL-1999 (Rel. 38, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hairless protein.
GN HR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

ALIGNMENTS

34	159	3.0	1745	1	CA35_HUMAN	P25940 homo sapien
35	159	3.0	2003	1	NTC4_HUMAN	Q99466 homo sapien
36	158	2.9	839	1	GLT5_WHEAT	P10388 triticum ae
37	158	2.9	1763	1	CA24_ASCSU	P27393 ascaris suu
38	156.5	2.9	1300	1	SAL3_HUMAN	Q9bxa9 homo sapien
39	156.5	2.9	1687	1	Z142_HUMAN	P52746 homo sapien
40	156	2.9	1106	1	GLI1_HUMAN	P08151 homo sapien
41	156	2.9	5120	1	PCLO_CHICK	Q9pu36 gallus gall
42	155.5	2.9	1364	1	CA21_BOVIN	P02465 bos taurus
43	155.5	2.9	1493	1	M3K1_RAT	Q62925 rattus norv
44	154	2.9	2441	1	CBP_MOUSE	P45481 mus musculu
45	153.5	2.9	1464	1	CA1I_HUMAN	P02452 homo sapien

Db 326 CCSSYPPTKGGDLPGCKQCEGLEGGASGASEPSEEVNKGASGRACPPSHHTTKLKTWLT 385

QY 181 RHSEQFECPRGCPVEEVRPVARLRALKRAGSPEVQAGMGSPAPKRPPDPFPGTAEQAGG 240

Db 386 RHSEQFECPRGCPVEEVRPVARLRALKRAGSPEVQAGMGSPAPKRPPDPFPGTAEQAGG 445

QY 241 WQEVRTDSIGNKDVDSQGHDEQKGPQDQASLQDPLQIPCLALPAKLAQCQSCAQAG 300

Db 446 WQEVRTDSIGNKDVDSQGHDEQKGPQDQASLQDPLQIPCLALPAKLAQCQSCAQAG 505

QY 301 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSGDSRLSTGLAKHLLSLGLGDR 360

Db 506 EGGHACHSQVRRSPGLGELQOEEDTATNSSSEEGPGSGDSRLSTGLAKHLLSLGLGDR 565

QY 361 LCRLLRREREALAWAQREGQGAFTGDSPPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420

Db 566 LCRLLRREREALAWAQREGQGAFTGDSPPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 625

QY 421 GRVACTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWKFD 480

Db 626 GRVACTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWKFD 685

QY 481 IRGHCPCQADARVWAPGDAGQKESTQKTPTTPQPSGNDTHRTKSIKEETPDSAETPAE 540

Db 686 IRGHCPCQADARVWAPGDAGQKESTQKTPTTPQPSGNDTHRTKSIKEETPDSAETPAE 745

QY 541 DRAGRGPLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDRITNILDSTIAQVV 600

Db 746 DRAGRGPLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDRITNILDSTIAQVV 805

QY 601 ERKTQEKALGPGLRAGPGLRGLPLSPVRPRLPPPGALLWLQEPQPCRRRGFHLFQEH 660

Db 806 ERKTQEKALGPGLRAGPGLRGLPLSPVRPRLPPPGALLWLQEPQPCRRRGFHLFQEH 865

QY 661 WRQGPVLVSGIQTTLQNLWGTEALGALGGQVQALSPLGPPQPSLSTTFWEGFSWPE 720

Db 866 WRQGPVLVSGIQTTLQNLWGTEALGALGGQVQALSPLGPPQPSLSTTFWEGFSWPE 925

QY 721 LRPKSDGVSLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 780

Db 926 LRPKSDGVSLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 985

QY 781 PQLWAAAYGVSPHRGLGTKNLCVEVADLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840

Db 986 PQLWAAAYGVSPHRGLGTKNLCVEVADLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 1045

QY 841 PGSQVSTVWHVFRQAQRIIRFLQMVCPAGAGALEPGAGSCYLDAGLRRRLREENGVS 900

Db 1046 PGSQVSTVWHVFRQAQRIIRFLQMVCPAGAGALEPGAGSCYLDAGLRRRLREENGVS 1105

QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFLSPETSALSQALCHQGPSLPPDCHL 960

Db 1106 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFLSPETSALSQALCHQGPSLPPDCHL 1165

QY 961 LYAQMDWAVFQAVKVAVGTLOEAK 984

Db 1166 LYAQMDWAVFQAVKVAVGTLOEAK 1189

RESULT 2

HAIR_MOUSE STANDARD; PRT; 1182 AA.

AC Q61645; Q80V47;

DT 01-NOV-1997 (Rel. 35, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hairless protein.

GN HR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c; TISSUE=Skin;

RX MEDLINE=94329587; PubMed=8052649;

RA Cachon-Gonzalez M.B., Fenner S., Coffin J.M., Moran C., Best S.,

RA Stoye J.P.;

RT "Structure and expression of the hairless gene of mice.";

RL Proc. Natl. Acad. Sci. U.S.A. 91:7717-7721(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Retina;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- FUNCTION: May act as a transcription factor that could act on to

CC regulate one of the phases of hair growth.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- TISSUE SPECIFICITY: Expressed predominantly in brain, hair

CC follicles and interfollicular epidermis. No expression in

CC dermis.

CC -!- DISEASE: HR MUTATION PRODUCES A NUMBER OF PLEIOTROPIC EFFECTS

CC INCLUDING STRUCTURAL ABNORMALITIES OF EPITHELIAL CELLS IN THE HAIR

CC FOLLICLES, HAIR LOSS TOWARDS THE END OF THE FIRST HAIR GROWTH

CC CYCLE, AND THE FAILURE OF SUBSEQUENT HAIR GROWTH CYCLES. OLDER

CC MICE CARRYING AN HR MUTATION HAVE BEEN REPORTED TO POSSESS ALTERED

CC RATIOS OF T-CELL-DEPENDENT B-CELL RESPONSES. MICE HOMOZYGOUS FOR

CC HR MUTATION ARE UNIQUELY SENSITIVE TO UV AND CHEMICALLY INDUCED

CC SKIN TUMORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z32675; CAA83587.1; --

EMBL; BC049182; AAH49182.1; --

PIR; I48378; I48378.

MGD; MGI:96223; hr.

InterPro; IPR003347; TF_JmjC.

Pfam; PF02373; jmjC; 1.

SMART; SM00558; JmjC; 1.

Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation; Metal-binding.

ZN_FING 595 620 C6-TYPE.

DOMAIN 535 540 POLY-GLY.

CONFLICT 401 401 P -> S (IN REF. 1).

SEQUENCE 1182 AA; 127192 MW; 3AFABE96C6EB3241 CRC64;

Query Match 76.7%; Score 4110.5; DB 1; Length 1182;

Best Local Similarity 78.5%; Pred. No. 3.5e-218;

Matches 772; Conservative 62; Mismatches 143; Indels 7; Gaps 5;

QY 1 FYKDPSPRLAKEPLAAABPGLFGLNSGGHLQRAEABRPSLHQDGMGAGRQNP 60

Db 206 FYHKDPNLLRPAKEPL--AESGMLGLAPGGHLQQAACESEGPSLHQDGETGAGRQQLCP 263
QY 61 LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPDGNLGYQLGPPATPRCPSPPEPVTQRG 120
Db 264 VFLGYDTPVPRAPWPCPPGLVHSLGNVWAGPSNSLGYQLGPPATPRCPSPGPPTPPGG 323
QY 121 CCSSYPPTKGGDLGPCGKCGQEGGASGASEPSEEVNKASGRACPPSHHTKLKKTWLT 180
Db 324 CCSSHLPAEGDGLGPCRKQDSPEGSSGPGSESSEERNKADS-RACPPSHHTKLKKTWLT 382
QY 181 RHSEQFECPRGCPVEEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFPGTAEQGAGG 240
Db 383 RHSEQFECPGGCGKKEESPATGLRALKRAGSPEVQASRGAPAPKRPSHPFPGTGRQGARA 442
QY 241 WQEVDRDTSIGNKDVDSGQHQDEKQPGDQASLQDPLQDIPCLALPAKLAQCQSCAAAG 300
Db 443 WQETPETIGSK-AEAQQEEQGRGPRDGRIRLQESRLVDTSCQHLAGVTQCQSCVQAAG 501
QY 301 EGGGHACHSQVRRSPLGGELQEEEDTATNSSSEEGSGPDSRLSTGLAKHLLSGLGDR 360
Db 502 EVGLTGHQSKRRSPL-EEKQLEEDSDSATSEEGGGPGPEASLNKGLAKHLLSGLGDR 560
QY 361 LCRLRRERREALAWAQREGQGPVAVTGPSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 561 LCRLLRERREALAWAQREGQGPAMTSDSPGIPHCCSRCHHGLFNTHWRCSHHRLCVAC 620
QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVVKFD 480
Db 621 GRIAGAGKNREKTGSQEHTDDCAQEAAGAACSLILTQFVSSQALAEELSTMVHQVWAKFD 680
QY 481 IRGHCPQADARVWAPGDAGQKQESTQKTPPTPQPSGNDTHRTKSIKEETPDSETPAE 540
Db 681 IRGHCFQVDARVWAPGDGGQKQKTEKTPPTPQPSGNDSTNRDKIKEETPDSETPAE 740
QY 541 DRAGRGLPCPSLCELLASTAVKLCGLHERIHMAFAPVTPALPSDDRITNILDIIAQVV 600
Db 741 DGAGRSPLPCPSLCELLASTAVKLCGLHRIHMAFAPVTPALPSDDRITNILDIIAQVV 800
QY 601 ERKIQEALGPGLRAGGLRKGLPLSPVPRPRLPPGALLWLQEPQPCRRGFHLFQEH 660
Db 801 ERKIQEALGPGLRAGGLRKGLSLPLSPVTRLSPPGALLWLQEP--PKHGFHLFQEH 858
QY 661 WRQGPVLVSGIQTILQNLWGTEALGALGGQVQALSPLGPPQPSISGTTFWEGFSWPE 720
Db 859 WRQGPVLVSGIQTILRLSLWGMEALGTGGQVQTLTALGPPQPTNLDSTAFWEGFSHPE 918
QY 721 LRPKSDGCVLLHHRALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 780
Db 919 TRPKLDEGCVLLHHRTLGDKOASRVQNLASSLPLPEYCAHQGKLNLAASYLPLGLTLHPLE 978
QY 781 POLMAAYGVSPHRGHLGTKNLCVEVADLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 979 POLMAAYGVNSHRGHLGTKNLCVEVSDLSILVHAESAQLPPWYRAQKDFLSGLDGEGLWS 1038
QY 841 PGQVSTVWHVFRAQDAQRIIRFLQVCPAGAGALEPGAGSCYLDAGLRRRLREEWGVS 900
Db 1039 PGQVSTVWHVFRAQDAQRIIRFLQVCPAGAGTLEPGAGSCYLDAGLRRRLREEWGVS 1098
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFVLSPELSALSQALCHQGPSLPPDCHL 960
Db 1099 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFVLSPELSALSQALCHQGPSLPPDHRM 1158
QY 961 LYAQMMDVAFQAVKAVGTLQEA 984
Db 1159 LYAQMMDRAVFAVKAAGALQEA 1182

RESULT 3

HAIR RAT

ID HAIR RAT

AC P97609;

DT 15-JUL-1998 (Rel. 36, Created)

STANDARD; PRT; 1181 AA.

DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Hairless protein.
GN HR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley;
RX MEDLINE=97141510; PubMed=8987811;
RA Thompson C.C.;
RT "Thyroid hormone-responsive genes in developing cerebellum include a
RT novel synaptotagmin and a hairless homolog.";
RL J. Neurosci. 16:7832-7840(1996).
CC -!- FUNCTION: May act as a transcription factor that could act on to
CC regulate one of the phases of hair growth.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U71293; AAC53018.1; ALT_INIT.
DR InterPro; IPR003347; TF_JmJC.
DR Pfam; PF02373; JmJC; 1.
DR SMART; SM00558; JmJC; 1.
KW Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;
KW Metal-binding.
FT ZN FING 594 619 C6-TYPE.
SQ SEQUENCE 1181 AA; 127307 MW; 834H7029CF8E88F0 CRC64;

Query Match

Best Local Similarity 75.9%; Score 4067; DB 1; Length 1181;

Matches 765; Conservative 65; Mismatches 146; Indels 8; Gaps 5;

QY 1 FYYKDPISIPRLAKEPLAAEPGLFGLNSGGHLQRAAGEAERPSLHQDGEWAGRQQLNCP 60
Db 206 FYHKDPNLLRPAKEPLAAESGMLGLAPGGHLQQAACESEGPSLHQDGETGAGRQQLCP 265
QY 61 LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPDGNLGYQLGPPATPRCPSPPEPVTQRG 120
Db 266 VFLGYDTPVPRTPWPCPPGLVHTLGNVWAGPSNSFGYQLGPPVTRCPSPGPPTPPGG 325
QY 121 CCSSYPPTKGGDLGPCGKCGQEGGASGASEPSEEVNKASGRACPPSHHTKLKKTWLT 180
Db 326 CCSSHLPAEGDGLGPCRKQDSPEGSSGPGSESSEERNKA-GSRASPPSHHTKLKKTWLT 384
QY 181 RHSEQFECPRGCPVEEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFPGTAEQGAGG 240
Db 385 RHSEQFECPGGCGKKEESPATGLRALKRAGSPEVQGARG-PAPKRPSHTFPGTGRQGARA 443
QY 241 WQEVDRDTSIGNKDVDSGQHQDEKQPGDQASLQDPLQDIPCLALPAKLAQCQSCAAAG 300
Db 444 WQETPETSTGSK-AEAQQEEQGRGPRDGRIRLRESRLDTSCQHLAGVTQCQSCVQAAG 502
QY 301 EGGGHACHSQVRRSPLGGELQEEEDTATNSSSEEGSGPDSRLSTGLAKHLLSGLGDR 360
Db 503 EVEILTSHSQSKHKLPLEEKPLEEDSCAT--SEEGGSSPEASINKGLAKHLLSGLGDR 559
QY 361 LCRLRRERREALAWAQREGQGPVAVTGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 560 LCRLLRERREALAWAQREGQGPAMTSDSPGIPHCCSRCHHGLFNTHWRCSHHRLCVAC 619
QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVVKFD 480
Db 620 GRIAGAGKNREKTGSREQRTDDCAQEAAGAACSLILTQFVSSQALAEELSTMVHQVWAKFD 679

QY 878 GAPGSCYLDAGLRRLREEWGVCWTLQAPGEAVLPAGAPHQVQGLVSTVSTQHFLS 937
Db 1093 IHDQSWYLDRLSLRKRLYQYGVQWAIQVFLGDDVVFIPAGAPHQVHNLISCIKVAEDFVS 1152
QY 938 PETA-----LSAQLCHQGPSLPPDCHLLYAQMDWAVFQAVKAVGTIQ 981
Db 1153 PEHVKCFWLQEFRLH--SQHTNHEDKLQVKNVIYHAVKDAVAMLK 1198

RESULT 5
APXL HUMAN
ID APXL HUMAN STANDARD; PRT; 1616 AA.
AC Q13796;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Apical-like protein (APXL protein).
GN APXL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=95315933; PubMed=77955590;
RA Schiaffino V.M., Bassi M.T., Rugarli E.I., Renieri A., Galli L.,
RA Ballabio A.;
RT "Cloning of a human homologue of the Xenopus laevis APX gene from the
RT ocular albinism type 1 critical region.";
RL Hum. Mol. Genet. 4:373-382(1995).
RN [2]
RP SEQUENCE OF 56-1616 FROM N.A.
RA Shen Y., Gibbs R.A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- TISSUE SPECIFICITY: Abundant in retina and melanoma; also in
CC brain, placenta, lung, kidney and pancreas.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -!- SIMILARITY: SOME, TO XENOPUS LAEVIS APICAL PROTEIN (APX).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X83543; CAA58534.1; -.
DR EMBL; AC002365; AAC32592.1; -.
DR PIR; I37183; I37183.
DR HSSP; P29476; 1QAV.
DR Genew; HGNC:630; APXL.
DR MIM; 300103; -.
DR GO; GO:0015280; F:amiloride-sensitive sodium channel activity; TAS.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS50106; PDZ; 1.
FT DOMAIN 26 108 PDZ.
FT DOMAIN 150 153 POLY-SER.
FT DOMAIN 314 320 POLY-PRO.
FT DOMAIN 343 346 POLY-ALA.
FT DOMAIN 1065 1068 POLY-PRO.
SQ SEQUENCE 1616 AA; 176409 MW; 752406B5BC0B60A2 CRC64;

Query Match 3.8%; Score 205; DB 1; Length 1616;
Best Local Similarity 21.7%; Pred. No. 0.00068;
Matches 210; Conservative 70; Mismatches 310; Indels 376; Gaps 54;
QY 9 PRLAKEPLAAAEPLFGLNSGG-----HL-----QRAGFAER--PSLHQRD----- 47
Db 363 PELTDRPWRSAHPGSLGKSGGPGCPQEAHADGSWPPSPKDGASSRLQASLSSSDVRFPOS 422

QY 48 -----GEMGAGRQONPCP---LFLGQPDITVFWTSWPACPPGLV 82
Db 423 PHSGRHPPLYSDHSPLCADSLGQEPGAASFQNDSPQVRGLSSCDQKLGSGWQGPSPCVQ 482
QY 83 HTL--GNVWAG--PGDGNLG--YQLGPPA---TPR--CPSPEPPVTQGCCSSYPPTKGG 131
Db 483 GDQAAQLWAGCWPSDTALGALESLEPPTVGQSPRHHLPOPEGPPDAR----- 530
QY 132 DLGPCGKCOEGLEGASGASEPSEEVNKASGPRACPPSHHTKLKTLWLRHSEQFECPRG 191
Db 531 ETGRCYPLDKGAEGCSAGAQEP-----RAS--RAEKASQRLAASITWADGESSRI-CPQE 583
QY 192 CP-----EVERPVARLRALKRAGSPEVQAMGSP-----APKRPPDPF----- 230
Db 584 TPLLHSLTQEGKRRP-----ESSPEDSATRPPPFDAHVGPTRRSDFATTLRNEI 634
QY 231 -----PGTAEQAGGW-----QE-----VRD 246
Db 635 QMHRAKLQKSRSTVALTAAGEAEDGTGRWRAGLGGTQEGPLAGTYKDHLEAQAARVLA 694
QY 247 TSIGNKDVDSGQHDEQKGPQDQASLQDPGLQDIPCLALPAKLAQCQSCAQAGE-----G 302
Db 695 TSFKRRDLDPNPGDLY--PESLEHRMGDP--DTVPHF--WEAGLAQPPSSSTGGPHPPRIG 749
QY 303 GGHACHSQVRRSPLGGELQOEEDTATNSSSEEGPGSGPDRLSTGLAKHLLSGLGDRLC 362
Db 750 GRRRFTABQKLKSYSEPEKMNVEGLTRGYSFHQHPRTSED-----VGTAD 796
QY 363 RLLRREREAALAWAQREGQGPVATG---DSPGIPRCCSRCHHGL-----FNTHW 407
Db 797 RWKFFETSKVPQORPAQKQALHGIPRDKPERPRTAGTCEGTEPWSRTTSLGDSLNAH- 855
QY 408 RCPRCSHRLCVACGRVAGTGR-----AREKAGFQEQSAEECTQEAG--HAACSLM-- 455
Db 856 -----SAAEKAGTSDLPRRLGTFAEYQASKEQKPLEARSSGRCHSADDILDV 904
QY 456 -----LTQFVSSQALAEALSTAMHQVWVKFDIRGHCPQCADARVWAPGDAGQKEST--- 506
Db 905 SLDPQERPHVHGRSSSPSTDHYKQEAASVELRQA-----GDPGEPEELPSA 953
QY 507 -----QKTPPTPOPSCNGDTHRTKSIKEETPDSAETPAEDRAGRGLPCPSLCELLAST 560
Db 954 VRAEQGSTPRQADAQC-----REGSPGQQHPPSQKA---PNP-PTFSEL----- 995
QY 561 AVKLCGLGHERIHMAFAPVTP-----ALPSDDRITNILDIIAQVVERKIQEKALGPG 612
Db 996 -----SHCRGAPELPREGRAGTLPRDYRYSE--ESTPAD-----LGPR 1033
QY 613 LRAGPG---LRKGLGLPLSPV-----RPRLPPPALLWLQEPQPCPRRGFHLFQEHWR 662
Db 1034 AQS--PGSPLHARGQDSWPVSSALLSKRPAPQORPPP-----PKREPRR-----YR 1076
QY 663 --QGQPVLVSGIORTLQGNLWGTALGALGGVQVQALSPLGPPQPSL-----GS 709
Db 1077 ATDGAP-----ADAPVGVLG-----RPFPPTSPASLDVYVARLSLSHSP 1115
QY 710 TTFWEGFSWPELTPKS---DEGSVLLHLRALGDEDTSRVENLAASLPLPEYCALHGKLN 766
Db 1116 SVF--SSAQPDQTPKATVCERGS---QHVSGD-----ASRPLPE----- 1149
QY 767 ASYLPP 772
Db 1150 -ALLPP 1154

RESULT 6
YHL1_EBV
ID YHL1_EBV STANDARD; PRT; 660 AA.
AC P03181;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OMnucleic - nucleic search, using sw model
Run on: June 29, 2004, 11:48:34 ; Search time 12295 Seconds
(without alignments)
11287.873 Million cell updates/sec

Title: US-10-024-368-1
Perfect score: 3202
Sequence: 1 ttttactacaagatccgag.....tgtgcccttttatgtacagg 3202

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3202	100.0	3202	6	AR194210	AR194210 Sequence
2	3185	99.5	5709	9	AF039196	AF039196 Homo sapi
3	3042.4	95.0	3719	9	HSA277165	AJ277165 Homo sapi
4	2817.4	88.0	3564	9	AF361864	AF361864 Macaca mu
5	2308.4	72.1	3210	9	BC008946	BC008946 Homo sapi
6	1985.4	62.0	5599	10	BC049182	BC049182 Mus muscu
7	1970.8	61.5	5186	10	RNU71293	U71293 Rattus norv
8	1970	61.5	4017	10	MMHAIPRO	Z32675 M.musculus
9	1963.6	61.3	4017	10	AY312855	AY312855 Mus muscu
10	943.4	29.5	2128	4	AY130969	AY130969 Ovis arie
11	828	25.9	1027	4	AY279972	AY279972 Sus scrof
12	786.8	24.6	218931	9	AC105206	AC105206 Homo sapi
13	786.2	24.6	958	9	HSA277250	AJ277250 Homo sapi
14	595.8	18.6	2394	6	AX835172	AX835172 Sequence
15	595.8	18.6	2394	9	AK098053	AK098053 Homo sapi
16	533.2	16.7	538	6	AX677839	AX677839 Sequence
17	460.2	14.4	207136	10	AC122268	AC122268 Mus muscu
C 18	460.2	14.4	219288	2	AC131321	AC131321 Mus muscu
C 19	449.4	14.0	255506	2	AC097410	AC097410 Rattus no
20	338.8	10.6	45185	2	AC011166	AC011166 Homo sapi
21	306	9.6	3357	9	HSM804566	AL833253 Homo sapi
22	279.8	8.7	231632	2	AC119127	AC119127 Rattus no
23	258.6	8.1	43402	2	AC105207	AC105207 Homo sapi
24	243	7.6	448	9	HSA400829	AJ400829 Homo sapi
25	197.8	6.2	582	9	HSA277252	AJ277252 Homo sapi
26	177.6	5.5	345	9	HSA400837	AJ400837 Homo sapi
27	171.4	5.4	454	9	HSA400835	AJ400835 Homo sapi
28	171.4	5.4	43402	2	AC105207	AC105207 Homo sapi
29	169.4	5.3	542	9	HSA400830	AJ400830 Homo sapi
30	168	5.2	691	9	HSA277253	AJ277253 Homo sapi
31	166	5.2	458	9	HSA400828	AJ400828 Homo sapi
32	150.2	4.7	376	9	HSA277251	AJ277251 Homo sapi
33	140.2	4.4	410	9	HSA400832	AJ400832 Homo sapi
34	133.4	4.2	464	9	HSA400836	AJ400836 Homo sapi
C 35	133.4	4.2	45185	2	AC011166	AC011166 Homo sapi
36	122.2	3.8	516	9	HSA400834	AJ400834 Homo sapi
37	121.2	3.8	367	9	HSA400833	AJ400833 Homo sapi
38	117.8	3.7	305	9	HSA400826	AJ400826 Homo sapi
39	103.6	3.2	419	10	MMMS39	AJ002245 Mus muscu
40	98.4	3.1	309	9	HSA400825	AJ400825 Homo sapi
41	86.6	2.7	125020	9	AF429315	AF429315 Homo sapi
42	85.8	2.7	301	11	BV014248	BV014248 S212P6107
43	85.8	2.7	490	9	HSA400827	AJ400827 Homo sapi
C 44	76	2.4	125020	9	AF429315	AF429315 Homo sapi
45	73.2	2.3	376	9	HSA400831	AJ400831 Homo sapi

ALIGNMENTS

RESULT 1
AR194210
LOCUS AR194210
DEFINITION Sequence 1 from patent US 6348348.
ACCESSION AR194210
VERSION AR194210.1 GI:20240802
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3202)
AUTHORS Thompson,C.C.
TITLE Human hairless gene and protein
JOURNAL Patent: US 6348348-A 1 19-FEB-2002;
FEATURES Location/Qualifiers

linear PAT 20-APR-2002

Db 2041 TGGGGACAGAAAGCTCTTGGGGCACTTGAGGCCAGGTGCAGGCGCTGAGCCCCCTCGGA 2100

Qy 2101 CCTCCCAGCCAGCAGCCTGGCAGCACAAATTTCTGGAGGGCTTCTCCTGGCCTGAG 2160

Db 2101 CCTCCCAGCCAGCAGCCTGGCAGCACAAATTTCTGGAGGGCTTCTCCTGGCCTGAG 2160

Qy 2161 CTTCGCCCAAAGTCAGACGAGGCTCTGTCTCTCTGCTGCAACCGAGCTTTGGGGGATGAG 2220

Db 2161 CTTCGCCCAAAGTCAGACGAGGCTCTGTCTCTCTGCTGCAACCGAGCTTTGGGGGATGAG 2220

Qy 2221 GACACAGCAGGGTGGAGAACCTTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCCTC 2280

Db 2221 GACACAGCAGGGTGGAGAACCTTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCCTC 2280

Qy 2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCTTGCCTCGCTCCACTGGAG 2340

Db 2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCTTGCCTCGCTCCACTGGAG 2340

Qy 2341 CCCAGCTCTGGGCAGCCTATGCTGAGCCCGCACCGGGGACACCTGGGACCAAGAAC 2400

Db 2341 CCCAGCTCTGGGCAGCCTATGCTGAGCCCGCACCGGGGACACCTGGGACCAAGAAC 2400

Qy 2401 CTCTGTGTGGAGGTGGCCGACCTGCTCAGCATCTGTGTGATGCGGACACACCACTGCCT 2460

Db 2401 CTCTGTGTGGAGGTGGCCGACCTGCTCAGCATCTGTGTGATGCGGACACACCACTGCCT 2460

Qy 2461 GCCTGGCACCGGGCACAGAAAGACTTCTTTTTCAGGCCTGGACGGGGAGGGCTCTGGTCT 2520

Db 2461 GCCTGGCACCGGGCACAGAAAGACTTCTTTTTCAGGCCTGGACGGGGAGGGCTCTGGTCT 2520

Qy 2521 CCGGGAGCCAGGTGAGCATGTGTGGACGTGTTCGGGGCACAGGACGCCAGCGCATC 2580

Db 2521 CCGGGAGCCAGGTGAGCATGTGTGGACGTGTTCGGGGCACAGGACGCCAGCGCATC 2580

Qy 2581 CGCCGCTTTCTCCAGATGGTGTGCCCGGCCGGGGCAGCGCCCTGGAGCCTGGCGCCCA 2640

Db 2581 CGCCGCTTTCTCCAGATGGTGTGCCCGGCCGGGGCAGCGCCCTGGAGCCTGGCGCCCA 2640

Qy 2641 GGCAGCTGTACTGATGATGAGGCTGCGCGCGCGCTGCGGGAGAGTGGGCGTGAGC 2700

Db 2641 GGCAGCTGTACTGATGATGAGGCTGCGCGCGCGCTGCGGGAGAGTGGGCGTGAGC 2700

Qy 2701 TGCTGGACCTGTCTCCAGGCCCCCGGAGAGCCGCTGTGTGTCTGCTGAGGGCTCCCCAC 2760

Db 2701 TGCTGGACCTGTCTCCAGGCCCCCGGAGAGCCGCTGTGTGTCTGCTGAGGGCTCCCCAC 2760

Qy 2761 CAGGTGACGGCCCTGCTGAGCACAGTCAAGCTCACTCAGCATCTTCTCTCCCTGAGACC 2820

Db 2761 CAGGTGACGGCCCTGCTGAGCACAGTCAAGCTCACTCAGCATCTTCTCTCCCTGAGACC 2820

Qy 2821 TCTGCCCTCTCTGCTCAGCTCTGCCACAGGGACCCAGCCTTCCCTGACTGCCACCTG 2880

Db 2821 TCTGCCCTCTCTGCTCAGCTCTGCCACAGGGACCCAGCCTTCCCTGACTGCCACCTG 2880

Qy 2881 CTTTATGCCAGATGGAGTGGCTGTGTTCAGAGCAGTGAAGTGGCGGTGGGACATTA 2940

Db 2881 CTTTATGCCAGATGGAGTGGCTGTGTTCAGAGCAGTGAAGTGGCGGTGGGACATTA 2940

Qy 2941 CAGGAGGCCAAATAGAGGATGCTAGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG 3000

Db 2941 CAGGAGGCCAAATAGAGGATGCTAGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG 3000

Qy 3001 GTGCTCAGCCAGGCACAACTTCAGCAGGGGATGGCGCTAGGGGACTTGGGATTTCTGG 3060

Db 3001 GTGCTCAGCCAGGCACAACTTCAGCAGGGGATGGCGCTAGGGGACTTGGGATTTCTGG 3060

Qy 3061 TCAACCCCAAGCAGCCTCTGGGACAGAGGGGCACTCTGTTCCCTCCCTCTTAAG 3120

Db 3061 TCAACCCCAAGCAGCCTCTGGGACAGAGGGGCACTCTGTTCCCTCCCTCTTAAG 3120

Qy 3121 CCAACAACCAAGTGCACCAAGCTCACACCTGCTCTCTCAGGCTGGCATCTCCCCCAC 3180

Db 3121 CCAACAACCAAGTGCACCAAGCTCACACCTGTCTCTCAGGCTGGCATCTCCCCCAC 3180

Qy 3181 CCTGTGCCCTTTTATGTACAGG 3202

Db 3181 CCTGTGCCCTTTTATGTACAGG 3202

RESULT 2

AF039196 5709 bp mRNA linear PRI 15-APR-2002

LOCUS Homo sapiens putative single zinc finger transcription factor

DEFINITION protein (hairless) mRNA, complete cds.

ACCESSION AF039196

VERSION AF039196.3 GI:20149786

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5709)

AUTHORS Ahmad, W., ul Haque, M.F., Brancolini, V., Tsou, H.C., ul Haque, S., Lam, H., Aita, V.M., Owen, J., deBlaguiera, M., Frank, J., Cserhalmi-Friedman, P.B., Leask, A., McGrath, J.A., Peacocke, M., Ahmad, M., Ott, J. and Christiano, A.M.

TITLE Alopecia universalis associated with a mutation in the human hairless gene

JOURNAL Science 279 (5351), 720-724 (1998)

MEDLINE 98111413

PUBMED 9445480

REFERENCE 2 (bases 1 to 5709)

AUTHORS Ahmad, W., Zlotogorski, A., Panteleyev, A.A., Lam, H., Ahmad, M., ul Haque, M.F., Abdallah, H.M., Dragan, L. and Christiano, A.M.

TITLE Genomic organization of the human hairless gene (HR) and identification of a mutation underlying congenital atrichia in an Arab Palestinian family

JOURNAL Genomics 56 (2), 141-148 (1999)

MEDLINE 99162400

PUBMED 10051399

REFERENCE 3 (bases 1 to 5709)

AUTHORS Christiano, A.M.

TITLE Direct Submission

JOURNAL Submitted (18-DEC-1997) Dermatology, Columbia University, 630 West 168th Street VC-1526, New York, NY 10032, USA

REFERENCE 4 (bases 1 to 5709)

AUTHORS Christiano, A.M.

TITLE Direct Submission

JOURNAL Submitted (18-AUG-1998) Dermatology, Columbia University, 630 West 168th Street VC-1526, New York, NY 10032, USA

REMARK Sequence update by submitter

REFERENCE 5 (bases 1 to 5709)

AUTHORS Christiano, A.M.

TITLE Direct Submission

JOURNAL Submitted (11-JUN-1999) Dermatology, Columbia University, 630 West 168th Street VC-1526, New York, NY 10032, USA

REMARK Sequence update by submitter

REFERENCE 6 (bases 1 to 5709)

AUTHORS Christiano, A.M.

TITLE Direct Submission

JOURNAL Submitted (15-APR-2002) Dermatology, Columbia University, 630 West 168th Street VC-1526, New York, NY 10032, USA

REMARK Sequence update by submitter

COMMENT On Apr 15, 2002 this sequence version replaced gi:5042382.

FEATURES

source

1. .5709

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="8"

/map="8p12"

1. .5709

/gene="hairless"

1485. .5054

/gene="hairless"

gene

CDS

/note="restricted expression in the brain and skin"
/codon_start=1
/product="putative single zinc finger transcription factor protein"
/protein_id="AAC32258.3"
/db_xref="GI:20149787"
/translation="MESTPSFLKGTPTWEKTAPENGIVRQEPGSPPRDGLHHGPLCLG
EPAPFWRGVLSTPDSWLPFPQGPQPKDMLPLVEGEGPQNGERKVNWLGSKEGLRWKEA
MLTHPLAFCGPACPPRCGLPMPEHSGHLKSDPVAFRPWCHPFLLETKILERAPFVWP
TCLPPYLVSGLPPEHPDCDWLTPHPWVYSGGPKVPSAFSLGSKGYKDPSPILRLAK
EPLAAEPLGLFNLSSGHLQORAGEAERPSLHQRDGEMAGRQNPCLFLGQPDTPVP
TSWPACPPGLVHTLGNVWAGPDGNLGYQLPPTAPRCPSPPEPVTQGCCSSYPPTK
GGDLGPCGKQEGLEGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLTRHSEQFE
CPRGCEVEERPVARLRALKRAGSPEVOQAMGSPAPKRPPDPFPPTAEQAGAGWQEV
DTSIGNKVDVSGQDEQKGPQDQASLQDPGLQDIPCLALPAKLAQCSQAQAAGEG
GHACHSQVRSPLGDELQDEEDATNSSEEGPGSGPDSRLSTGLAKHLLSGLGDR
CRLRRERREALAWAQREGQPAVTEDSPIPRCCSRCHHGLFNTHWRCPRCSHRLCVA
CGRVAGTGRAREKAGFQEQSAEECTQEAHAAACSLMLTQFVSSQALAEALSAMHQQVW
KFDIRHCHCQADARVWAPGDAGQKESTQKTPPTPQPSNGDTHRTKSIKEETPDSA
ETPAEDRAGRGPLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPSSDDRITNILD
SIIIAVVERKIQEKALGPLRAGPLRKLGLPLSPVRPRPLPPGALLWLQEPQPCPR
RGFHLFQHWKRGQPVLVSGIQRITLQNLWGTEALGALGVQVQALSPGPPQPSGLS
TTFWEGFSWPELRPKSDEGSVLLHRLALGDEDTSRVENLAASPLPEYCALHGLKSLA
SYLPPGLALRPLEPQLWAAVYGVSPHRGHGTLKNLCVEADLVSLVHADTPLPAWHRA
QKDFLSLDGEGLSWSPGSQVSTVWHVFERAQAQRIIRFLQMVCPAGAGALEPAGPSC
YLDAGLRRLREEWGVSCWTLTLOAPGEAVLVPAGAPHQVQGLVSTVSTQHFLSPETS
ALSAQLCHQGPSLPPDCHLLYAQMDWAVFQAVKVAVGTLQLEAK"

ORIGIN

Query Match		99.5%;	Score 3185;	DB 9;	Length 5709;
Best Local Similarity		99.8%;	Pred. No. 0;		
Matches 3188;		Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
QY	1	TTTTACTACAGGATCCGAGCAT	TCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA	60	
Db	2100	TTTTACTACAGGATCCGAGCAT	TCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA	2159	
QY	61	CCTGGGTTGTTGGCTTAAACT	CTGGTGGGACCTGCAGAGAGCCGGGAGGCCGAACGC	120	
Db	2160	CCTGGGTTGTTGGCTTAAACT	CTGGTGGGACCTGCAGAGAGCCGGGAGGCCGAACGC	2219	
QY	121	CCTTCACTGCACAGAGGGATG	GAGAGATGGAGCTGGCCGGCAGCAGAACTCTTGCCCG	180	
Db	2220	CCTTCACTGCACAGAGGGATG	GAGAGATGGAGCTGGCCGGCAGCAGAACTCTTGCCCG	2279	
QY	181	CTCTTCTGGGAGCCAGACACT	GTGCCCTTGACCTCTGGCCCGCTTGTCCTCCCGCAGGC	240	
Db	2280	CTCTTCTGGGAGCCAGACACT	GTGCCCTTGACCTCTGGCCCGCTTGTCCTCCCGCAGGC	2339	
QY	241	CTTGTTCATCTCTTGGCAAC	GTCTGGGCTGGCCAGGCGATGGGAACCTTGGGTACCA	300	
Db	2340	CTTGTTCATCTCTTGGCAAC	GTCTGGGCTGGCCAGGCGATGGGAACCTTGGGTACCA	2399	
QY	301	CTGGGGCCACAGCAACACCA	AGGTGCCCTCTCTGAGCCGCTGTCAACCCAGCGGGC	360	
Db	2400	CTGGGGCCACAGCAACACCA	AGGTGCCCTCTCTGAGCCGCTGTCAACCCAGCGGGC	2459	
QY	361	TGCTGTTCTATCTACCCACC	CTAAAGTGGGGATCTTGGCCCTTGTGGGAAGTGCCAG	420	
Db	2460	TGCTGTTCTATCTACCCACC	CTAAAGTGGGGATCTTGGCCCTTGTGGGAAGTGCCAG	2519	
QY	421	GAGGGCCTGGAGGGGGTGC	CCAGTGGAGCCAGCAACCCAGCGAGGAAGTGAACAAGGCC	480	
Db	2520	GAGGGCCTGGAGGGGGTGC	CCAGTGGAGCCAGCAACCCAGCGAGGAAGTGAACAAGGCC	2579	
QY	481	TCTGGCCCCAGGGCCTGT	TCCCCCAGCCACACCAAGCTGAAGAAGACATGGCTCACA	540	
Db	2580	TCTGGCCCCAGGGCCTGT	TCCCCCAGCCACACCAAGCTGAAGAAGACATGGCTCACA	2639	
QY	541	CGGCACCTCGGAGAGTTT	GAAATGTCCACGCGGCTGCCCTGAGGTCGAGGAGGCCGGTT	600	
Db	2640	CGGCACCTCGGAGAGTTT	GAAATGTCCACGCGGCTGCCCTGAGGTCGAGGAGGCCGGTT	2699	

QY	601	GCTCGGCTCCGGGCCCTCA	AAAGGGCAGGAGCCCGAGGTCCAGGAGCAATGGGCAGT	660	
Db	2700	GCTCGGCTCCGGGCCCTCA	AAAGGGCAGGAGCCCGAGGTCCAGGAGCAATGGGCAGT	2759	
QY	661	CCAGCCCCCAAGCGGCAC	CCGGACCCCTTTCCAGGCACTGCAGAACAGGGGCTGGGGT	720	
Db	2760	CCAGCCCCCAAGCGGCAC	CCGGACCCCTTTCCAGGCACTGCAGAACAGGGGCTGGGGT	2819	
QY	721	TGGCAGGAGGTTCCGGSA	CACATCGATAGGGAACAAGGATGTGACTCGGGACAGCATGAT	780	
Db	2820	TGGCAGGAGGTTCCGGSA	CACATCGATAGGGAACAAGGATGTGACTCGGGACAGCATGAT	2879	
QY	781	GAGCAGAAAGGACCCCA	AGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA	840	
Db	2880	GAGCAGAAAGGACCCCA	AGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA	2939	
QY	841	CCATGCCTGCTTCTCCCT	GCATAAATGGCTCAATGCCAAAAGTTGTCCCCAGGAGCTGGA	900	
Db	2940	CCATGCCTGCTTCTCCCT	GCATAAATGGCTCAATGCCAAAAGTTGTCCCCAGGAGCTGGA	2999	
QY	901	GAGGAGGAGGCGCGCTG	CCACTCTCAGCAAGTGCAGGATCGCCTCTGGGAGGGAG	960	
Db	3000	GAGGAGGAGGCGCGCTG	CCACTCTCAGCAAGTGCAGGATCGCCTCTGGGAGGGAG	3059	
QY	961	CTGCAGCAGGAGGAAGA	CACACAGCCCAACTCCAGCTCTGAGGAAGGCCAGGCTCCGGC	1020	
Db	3060	CTGCAGCAGGAGGAAGA	CACACAGCCCAACTCCAGCTCTGAGGAAGGCCAGGCTCCGGC	3119	
QY	1021	CCTGACAGCCGGCTCAG	CACAGGCCCTCGCCAAAGCACCTGCTCAGTGGTTTGGGGACCGA	1080	
Db	3120	CCTGACAGCCGGCTCAG	CACAGGCCCTCGCCAAAGCACCTGCTCAGTGGTTTGGGGACCGA	3179	
QY	1081	CTGTGCCGCTCTGTCGG	AGGAGCGGCTTGGCCCCAGCGGGGAAGGCCAA	1140	
Db	3180	CTGTGCCGCTCTGTCGG	AGGAGCGGCTTGGCCCCAGCGGGGAAGGCCAA	3239	
QY	1141	GGGCCAGCCGTGACAGG	GGACAGCCAGCTGTCAGCGCTGTCAGCGCTTGCACCAT	1200	
Db	3240	GGGCCAGCCGTGACAGG	GGACAGCCAGCTGTCAGCGCTGTCAGCGCTTGCACCAT	3299	
QY	1201	GGACTCTTCAACACCC	ACTGGCGATGTCCCGCTGCAGCCACCGGCTGTGTGGCCTGT	1260	
Db	3300	GGACTCTTCAACACCC	ACTGGCGATGTCCCGCTGCAGCCACCGGCTGTGTGGCCTGT	3359	
QY	1261	GGTCTGTGGCAGGCACT	GGCGGCGCCAGGAGAAAAGCAGGCTTTCAGGAGCAGTCCGCG	1320	
Db	3360	GGTCTGTGGCAGGCACT	GGCGGCGCCAGGAGAAAAGCAGGCTTTCAGGAGCAGTCCGCG	3419	
QY	1321	GAGGAGTGACGAGAGG	CCCGGCACGCTGCCTGTTCCTGATGCTGACCCAGTTTGTCT	1380	
Db	3420	GAGGAGTGACGAGAGG	CCCGGCACGCTGCCTGTTCCTGATGCTGACCCAGTTTGTCT	3479	
QY	1381	TCCAGCCAGGCTTTGG	CAGAGCTGAGCACTGCAATGCACCAAGTCTGGGTCAAGTTTGT	1440	
Db	3480	TCCAGCCAGGCTTTGG	CAGAGCTGAGCACTGCAATGCACCAAGTCTGGGTCAAGTTTGT	3539	
QY	1441	ATCCGGGGGACTGCCC	CTGCCAAGCTGATGCCCGGGTATGGGCCCCCGGGGATGCAGGC	1500	
Db	3540	ATCCGGGGGACTGCCC	CTGCCAAGCTGATGCCCGGGTATGGGCCCCCGGGGATGCAGGC	3599	
QY	1501	CAGCAGAAGGAATCAAC	ACAGAAAAACGCCCCAACTCCACAACCTTCTCTGCAATGGCGAC	1560	
Db	3600	CAGCAGAAGGAATCAAC	ACAGAAAAACGCCCCAACTCCACAACCTTCTCTGCAATGGCGAC	3659	
QY	1561	ACCCACAGGACCAAG	AGCATCAAAGAGGAGACCCCGGATTCGCTGAGACCCCGAGAG	1620	
Db	3660	ACCCACAGGACCAAG	AGCATCAAAGAGGAGACCCCGGATTCGCTGAGACCCCGAGAG	3719	
QY	1621	GACCGTGTGCGGAGG	CCCCCTGCCTTCTCTCTGCGAACTGCTGCTTCTTACC	1680	
Db	3720	GACCGTGTGCGGAGG	CCCCCTGCCTTCTCTCTGCGAACTGCTGCTTCTTACC	3779	
QY	1681	GCGGTCAAACACTCTG	CTTGGGGCCATGAGCGAATACATATGGCCCTTCGCCCCCGTCACTCCG	1740	


```
5'UTR
1. .40
CDS
/gene="HR"
/prot="HR"
41. .3610
/prot="HR"
/note="putative single zinc finger transcription factor
protein, responsible for autosomal recessive universal
congenital alopecia"
/codon_start=1
/product="hairless protein"
/protein_id="CAB86602.1"
/db_xref="GI:7529560"
/db_xref="GOA:O43593"
/db_xref="SWISS-PROT:O43593"
/translation="MESTPSFLKGTPTWEKTAPENGIVRQEPGSPPRDGLHHGPLCLG
EPAPFRWGLSTPDSWLPFGFQGPQKMDLPLVEGEGPQNGKRVNWLGSKEGLRWKEA
MLTHPLAFCPACPPRCGPLMPEHSGHLKSDPVAFRPWHCPLLETKILERAPFWVP
TCLPYLVSLPPEHPCDWPLTPHPWVYSGGQPKVPSAFSLGSKGFYKDPISIRLAK
EPLAAAEPLGLFGLNSGHLQRAGEAERP SLHQRDGEMAGRQONPCPLFLGQPDTPW
TSWPACPPGLVHTLGNVWAGDNLGYQLGPPATPRCPSPPEPPVTQRCSSSYRPTK
GGGLPGCGKQEGLEGASGASEPSEEVNKASGPRACPSPHHTKLKTLWLRHSEQFE
CPRGPEVEERPFVARLRALKRAGSPVEVQAMGSPAPKPPDPFPGTAEQAGAGWQEV
DTSIGNKDVDSGQHQDEKQPDQASLQDPGLQDI PCLALPAKLAQCQSCAAAGGG
GHACHQVRRSPLGGELQOEDTATNSSEEGPGSPDSRLSTGLAKHLLSGLGDL
CRLLREREALAWAQREGQPAVTEPSGIPRCCSRCHHGLFNTHWRCPRCSHRLCVA
CRVAGTGRAREKAGFEQSAEECTQEAHHAACMLTQFVSSQALAE LSTAMHQVW
KFDIRGHCPCQADARVWAPDAGQOKESTKTPPTPQSCNGDTHRTKSIKEETPPSA
ETPAEDRAGRGLPCPSLCELLASTAVKLC LGLHERIHMAFAPVTPALPSDDRIITILD
SIIAQVVERKIQEKALGPLRAGPLRKLGLPLSPVRPLPPPGALLWLQEPQPCPR
RGFHLFOEHWROGPVLVSGIQT LQGNLWTEALGALGGQVQALSPLGPPQPSLSGS
TTFWEGFSWPELRKSDSGVLLHRLALGDEDTSRVENLAASLPLPEYCALHGLKLNLA
SYLPPGLALRPLEPQLWAAVGSVPHRHLGTLKNLCVEVADLVSLVHADTPLPAWHRA
QKDFLSGLDGEGLWSPGSQVSTVWHVFRQAQDAQRIIRFLQWVCPAGAGALEPGAGSC
YLDAGLRRRLREEWGSCTLLQAPGEAVLVPAGAPHVQGLVSTVSTQHF LSPETS
ALSAQLCHQGPSLPPDCHLLYAQMDWAVFQAVKVAVGT LQEAK"
3611. .>3719
/prot="HR"
3'UTR
ORIGIN
```

```
Query Match 95.0%; Score 3042.4; DB 9; Length 3719;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3057; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 TTTTACTACAAGGATCCGAGCATTCCTCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA 60
Db |||||
656 TTTTACTACAAGGATCCGAGCATTCCTCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA 715
QY 61 CCTGGGTGTTTGGCTTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGGAACGC 120
Db |||||
716 CCTGGGTGTTTGGCTTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGGAACGC 775
QY 121 CCTTCACTGCACCAGAGGGATGGAGAGATGGGAGCTGGCCGGCAGCAGAAATCCTTGCCCG 180
Db |||||
776 CCTTCACTGCACCAGAGGGATGGAGAGATGGGAGCTGGCCGGCAGCAGAAATCCTTGCCCG 835
QY 181 CTCTTCTTGGGGCAGCCAGACACTGTGCCCTTGGACCTCTGGCCCTGGACCTCTGGCCCTGG 240
Db |||||
836 CTCTTCTTGGGGCAGCCAGACACTGTGCCCTTGGACCTCTGGCCCTGGACCTCTGGCCCTGG 895
QY 241 CTGTGTCATACTCTTGGCAACGCTGTGGGCTGGGCCAGGCATGGGAACCTTGGGTACCAG 300
Db |||||
896 CTGTGTCATACTCTTGGCAACGCTGTGGGCTGGGCCAGGCATGGGAACCTTGGGTACCAG 955
QY 301 CTGGGGCCACACAGCAACACCAAGTGCCCTCTCTCTGAGCCGCTGTCAACCCAGCGGGC 360
Db |||||
956 CTGGGGCCACACAGCAACACCAAGTGCCCTCTCTCTGAGCCGCTGTCAACCCAGCGGGC 1015
QY 361 TGCTGTTTCATCTACCCACCACTAAAGTGGGGATCTTGGCCCTTGTGGGAAGTGCAG 420
Db |||||
1016 TGCTGTTTCATCTACCCACCACTAAAGTGGGGTCTTGGCCCTTGTGGGAAGTGCAG 1075
QY 421 GAGGGCCTGGAGGGGTGTCAGTGGAGCCAGCAACCCAGCGAGGAAGTGAACAGGCC 480
Db |||||
```

```
Db 1076 GAGGGCCTGGAGGGGGTGCCAGTGGAGCCAGCGAAACCAGCGAGGAAGTGAACAAGGCC 1135
QY 481 TCTGGCCCCAGGGCCTGTCCCCCAGCCACACACCAAGCTGAAGAAGACATGGCTCACA 540
Db |||||
1136 TCTGGCCCCAGGGCCTGTCCCCCAGCCACACACCAAGCTGAAGAAGACATGGCTCACA 1195
QY 541 CGGCACCTCGGAGCAGTTTGAATGTCCACGGGCTGCCCTGAGGTCGAGGAGAGGCCGGTT 600
Db |||||
1196 CGGCACCTCGGAGCAGTTTGAATGTCCACGGGCTGCCCTGAGGTCGAGGAGAGGCCGGTT 1255
QY 601 GCTCGGCTCCGGGCCCTCAAAAGGGCAGGACGCCCGAGGTCAGGGAGCAATGGGCAGT 660
Db |||||
1256 GCTCGGCTCCGGGCCCTCAAAAGGGCAGGACGCCCGAGGTCAGGGAGCAATGGGCAGT 1315
QY 661 CCAGCCCCCAAGCGGCCACCGGACCCCTTTCCAGGCACCTGCAGAACAGGGGCTGGGGT 720
Db |||||
1316 CCAGCCCCCAAGCGGCCACCGGACCCCTTTCCAGGCACCTGCAGAACAGGGGCTGGGGT 1375
QY 721 TGGCAGGAGGTTCCGGGACACATCGATAGGGAACAAGGATGTGGACTCGGACACAGCATGAT 780
Db |||||
1376 TGGCAGGAGGTTCCGGGACACATCGATAGGGAACAAGGATGTGGACTCGGACACAGCATGAT 1435
QY 781 GAGCAGAAAGGACCCCAAGATGGCCAGGCAGTCTCCAGGACCCCGGACTTCAGGACATA 840
Db |||||
1436 GAGCAGAAAGGACCCCAAGATGGCCAGGCAGTCTCCAGGACCCCGGACTTCAGGACATA 1495
QY 841 CCATGCCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCAGGCAGCTGGA 900
Db |||||
1496 CCATGCCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCAGGCAGCTGGA 1555
QY 901 GAGGAGGAGGACCGCTGCCACTCTCAGCAAGTGCAGGATCGCCTCTGGGAGGGGAG 960
Db |||||
1556 GAGGAGGAGGACCGCTGCCACTCTCAGCAAGTGCAGGATCGCCTCTGGGAGGGGAG 1615
QY 961 CTGCAGCAGGAGGAAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCCAGGTCGCGC 1020
Db |||||
1616 CTGCAGCAGGAGGAAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCCAGGTCGCGC 1675
QY 1021 CCTGACAGCCGGCTCAGCACAGGCTCGCCAAAGCACCTGCTCAGTGGTTTGGGGACCGA 1080
Db |||||
1676 CCTGACAGCCGGCTCAGCACAGGCTCGCCAAAGCACCTGCTCAGTGGTTTGGGGACCGA 1735
QY 1081 CTGTGCCGCTGTCTCGGAGGGAGCGGAGGCCCTGGCTTGGGCCCAGCGGAAGGCCAA 1140
Db |||||
1736 CTGTGCCGCTGTCTCGGAGGGAGCGGAGGCCCTGGCTTGGGCCCAGCGGAAGGCCAA 1795
QY 1141 GGGCCAGCCGTGACAGGGGACAGCCAGGCATTTCCACGCTGCTGCAGCCGTTGCCACCAT 1200
Db |||||
1796 GGGCCAGCCGTGACAGAGGACAGCCAGGCATTTCCACGCTGCTGCAGCCGTTGCCACCAT 1855
QY 1201 GGACTCTTCAACACCCACTGGCGATGTCCCGTGCAGCCACCGCTGTGTGTGGCCTGT 1260
Db |||||
1856 GGACTCTTCAACACCCACTGGCGATGTCCCGTGCAGCCACCGCTGTGTGTGGCCTGT 1915
QY 1261 GGTGCTGTGGCAGGCACCTGGCGGGCCAGGGAGAAAAGCAGGCTTTCAGGAGCAGTCCGCG 1320
Db |||||
1916 GGTGCTGTGGCAGGCACCTGGCGGGCCAGGGAGAAAAGCAGGCTTTCAGGAGCAGTCCGCG 1975
QY 1321 GAGAGTGCACGCAGGAGGCGGGCAGCGTGCCTGCTTCCCTGATGTGACCCAGTTTGTGTC 1380
Db |||||
1976 GAGAGTGCACGCAGGAGGCGGGCAGCGTGCCTGCTTCCCTGATGTGACCCAGTTTGTGTC 2035
QY 1381 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACACAGGCTCTGGGTCAAGTTTGTAT 1440
Db |||||
2036 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACACAGGCTCTGGGTCAAGTTTGTAT 2095
QY 1441 ATCCGGGGGCACTGCCCTGCCAAGTGTATGCCCGGGTATGGGCCCCCGGGGATGCAGGC 1500
Db |||||
2096 ATCCGGGGGCACTGCCCTGCCAAGTGTATGCCCGGGTATGGGCCCCCGGGGATGCAGGC 2155
QY 1501 CAGCAGAAAGGAATCAACACAGAAAACGCCCCCAACTCCACAACTTCTCTGCAATGGCGAC 1560
Db |||||
2156 CAGCAGAAAGGAATCAACACAGAAAACGCCCCCAACTCCACAACTTCTCTGCAATGGCGAC 2215
```


QY	1561	ACCCACAGGACCAAGAGCATCAAAGAGGAGACCCCGATTCCGCTGAGACCCCGAGCAGAG	1621
Db	2216	ACCCACAGGACCAAGAGCATCAAAGAGGAGACCCCGATTCCGCTGAGACCCCGAGCAGAG	2275
QY	1621	GACCGTGCTGGCCGAGGGCCCCCTGCCCTTGTCCTTCTCTCTGCGAACTGCTGGCTTCTACC	1680
Db	2276	GACCGTGCTGGCCGAGGGCCCCCTGCCCTTGTCCTTCTCTCTGCGAACTGCTGGCTTCTACC	2335
QY	1681	GCGGTCAAACCTCTGCTTGGGCCATGAGCGAATACACATGCGCCTTCGCCCCCGTCACTCCG	1740
Db	2336	GCGGTCAAACCTCTGCTTGGGCCATGAGCGAATACACATGCGCCTTCGCCCCCGTCACTCCG	2395
QY	1741	GCCCTGCCCCAGTGATGACCGCATCACCAACATCCTCGACAGCATTATCGCACAGGTGGTG	1800
Db	2396	GCCCTGCCCCAGTGATGACCGCATCACCAACATCCTCGACAGCATTATCGCACAGGTGGTG	2455
QY	1801	GAACGGAGATCCAGGAGAAAGCCCTGGGGCCCGGGCTTCGAGCTGGCCCGGGTCTGCGC	1860
Db	2456	GAACGGAGATCCAGGAGAAAGCCCTGGGGCCCGGGCTTCGAGCTGGCCCGGGTCTGCGC	2515
QY	1861	AAGGGCCTGGGCCTGCCCTCTCTCCAGTCGGCCCGGGCTGCCTCCCCAGGGCTTTG	1920
Db	2516	AAGGGCCTGGGCCTGCCCTCTCTCCAGTCGGCCCGGGCTGCCTCCCCAGGGCTTTG	2575
QY	1921	CTGTGGCTGCAGGAGCCCCAGCCTTGCCCTCGGCGTGGCTTCACCTCTTCCAGGAGCAC	1980
Db	2576	CTGTGGCTGCAGGAGCCCCAGCCTTGCCCTCGGCGTGGCTTCACCTCTTCCAGGAGCAC	2635
QY	1981	TGAGGCAGGGCCAGCCTGTGTTGGTGTACAGGATCCAAAGGACATTGCAGGGCAACCTG	2040
Db	2636	TGAGGCAGGGCCAGCCTGTGTTGGTGTACAGGATCCAAAGGACATTGCAGGGCAACCTG	2695
QY	2041	TGGGGACAGAAGCTTTTGGGGCACTTGAGGGCCAGGTGCAGGCGTGAGCCCCCTCGGA	2100
Db	2696	TGGGGACAGAAGCTTTTGGGGCACTTGAGGGCCAGGTGCAGGCGTGAGCCCCCTCGGA	2755
QY	2101	CCTCCCCAGCCAGAGCCTGGGCAGCACAACTTCTGGGAGGGCTTCTCTGGCCTGAG	2160
Db	2756	CCTCCCCAGCCAGAGCCTGGGCAGCACAACTTCTGGGAGGGCTTCTCTGGCCTGAG	2815
QY	2161	CTTCGCCCAAAGTCAGACGAGGGCTCTGTCTCTGTCTGCACCGAGCTTTGGGGATGAG	2220
Db	2816	CTTCGCCCAAAGTCAGACGAGGGCTCTGTCTCTGTCTGCACCGAGCTTTGGGGATGAG	2875
QY	2221	GACACCAGCAGGTGGAGAACCTAGCTGCCAGTGTGCCACTTCGGAGTACTGCGCCCTC	2280
Db	2876	GACACCAGCAGGTGGAGAACCTAGCTGCCAGTGTGCCACTTCGGAGTACTGCGCCCTC	2935
QY	2281	CATGGAAAACTCAACTGGCTTCTACCTCCACCGGGCCTTGCCCTGCGTCCACTGGAG	2340
Db	2936	CATGGAAAACTCAACTGGCTTCTACCTCCACCGGGCCTTGCCCTGCGTCCACTGGAG	2995
QY	2341	CCCCAGCTCTGGGCAGCCTATGGTGTGAGCCCGCACCGGGACACCTGGGGACCAAGAAC	2400
Db	2996	CCCCAGCTCTGGGCAGCCTATGGTGTGAGCCCGCACCGGGACACCTGGGGACCAAGAAC	3055
QY	2401	CTCTGTGTGAGGTGGCCGACCTGGTCAGCATCTGTGTGCATGCCGACACACCACTGCCT	2460
Db	3056	CTCTGTGTGAGGTGGCCGACCTGGTCAGCATCTGTGTGCATGCCGACACACCACTGCCT	3115
QY	2461	GCCTGGCACCGGGCACAGAAAGACTTCCTTTACAGCCTGGACGGGAGGGGCTCTGGTCT	2520
Db	3116	GCCTGGCACCGGGCACAGAAAGACTTCCTTTACAGCCTGGACGGGAGGGGCTCTGGTCT	3175
QY	2521	CCGGGCAGCCAGGTACGACTGTGTGGCACGTGTTCGGGCACAGGACGCCAGCGCATC	2580
Db	3176	CCGGGCAGCCAGGTACGACTGTGTGGCACGTGTTCGGGCACAGGACGCCAGCGCATC	3235
QY	2581	CGCCGCTTTCTCCAGATGGTGTGCCCGGCCGGGACGGCCCTTGGAGCTTGGCGCCCCA	2640
Db	3236	CGCCGCTTTCTCCAGATGGTGTGCCCGGCCGGGACGGCCCTTGGAGCTTGGCGCCCCA	3295

```

RESULT 4
AF361864
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

```


GLPGCKQCEGLDGVASGASEPSEEVNKASGPRAcppGHHTKLKKTWLTRHSEQFECp
RGCPAEERpVqLRALKRAGSPEIQGAVGGpAPKRPpDPFGTAEOqAGGwQwEVrDT
SIGNKEADSgQhDDQRpRDGQASLQDPGLQIPLCLAPKLAQCQSCAQAAGEGGp
AGHFQqVRRSpLGGELQOEEDTAANSSSEEGpSGpDGRlSTGLSKHLLSSlGDRlCR
LLRREREALAwAQEGQpPAVTEDNPGIPRCCSRCHHGLFNTHWRCPRSHRlCvACG
RVAGARAREKAgSREqSTEECTOEAGHAACSILTLTQVSSQALAElSTAMHQVwVKF
DIRGHPCQADARVWAPGDAGQKESQKTPPTPQPScNGDTHTRTKSikEETPDSTET
PAEDRAGApLpCpSLCELLASTAVKLCLGHERIHMAFAPVTPALPSDDRITNlLDSI
IAQVVERKIqEKALGPGLRAGpGLRKLGLPLSPVRpRLPPGALLWlQRPpRlPQRG
FHLfQEHWRQGPVLVSGIQRTLQGNLWGTEALGALGGQVQALSPLPPQpOTSLSGSTA
FWEGfSWPElRPKSDegSVLLHlRALGDEdTSVENLAASLPLPEYCAHRGKlNLASy
LPPGLALRPLepOLwAAyGVSPhRGHLGTKNLCVEADLVSlLVHADTPlPAWHRAQK
DFLSGLDGEGLWSPGSQVSTVWHVFRaQDAQRIRrFLQMVCPAGAGALEPAGpSCYL
DAGLRRlRlREEWGVCWTLlQAPGEAVLVPAGAPHQVQGLVSTVSTQHFLSPETSAL
SAQLCHQGPSLpPDChLLLYAQMDWAVFQAVKVAVGTlQEAK"

ORIGIN

Query Match 88.0%; Score 2817.4; DB 9; Length 3564;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 2869; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY	1	TTTTACTACAAGGATCCGAGCATTTCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA	60
Db	610	TTTTACTACAAGGATCCGAGCATTTCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA	669
QY	61	CCTGGGTTGTTTGGCTTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGAACGC	120
Db	670	CCTGGGTTGTTTGGCTTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGAACGC	729
QY	121	CCTTCACTGCACCAGAGGGATGGAGAGATGGGAGCTGGCCGGCAGCAGAAATCCTTGCCCG	180
Db	730	CCTTCACTGCACCAGAGGGATGGAGAGATGGGAGCTGGCCGGCAGCAGAAATCCTTGCCCG	789
QY	181	CTCTTCCTGGGGCAGCCAGACACTGTGCCCTTGGACCTCCTGGCCCGCTTGTCCCCAGGC	240
Db	790	CTCTTCCTGGGGCAGCCAGACACTGTGCCCTTGGACCTCCTGGCCCGCTTGTCCCCAGGC	849
QY	241	CTTGTTTCATACTCTTGGCAAACGTCTGGGCTGGGCCAGGCGATGGAACCTTGGGTACCAG	300
Db	850	CTTGTTTCATACTCTTGGCAAACGTCTGGGCTGGGCCAAGCGGTGGAGCCTTGGGTACCAG	909
QY	301	CTGGGGCCACCAGCAACACCAAGGTGCCCCCTCTCCTGAGCCGCCCTGTCACCCAGCGGGGC	360
Db	910	CTGGGGCCACCACCAACCAAGGTGCCCCCTCTCCGAGCCGCCCTATCACCCAGCGGGGC	969
QY	361	TGCTGTTTCATCCTTACCCACCCATTAAGATGGGGTCTTGGCCCTTGTGGGAAGTGCCAG	420
Db	970	TGCTGTTTCATCCTTACCCACCCATTAAGATGGGGTCTTGGCCCTTGTGGGAAGTGCCAG	1029
QY	421	GAGGCGCTGGAGGGGGTGCAGTGGAGCCAGCGAAACCCAGCGAGGAAGTGAACAAGGCC	480
Db	1030	GAGGCGCTGGAGCGGGTGTCCAGTGGAGCCAGCGAAACCCAGCGAGGAAGTGAACAAGGCC	1089
QY	481	TCTGGCCCAAGGGCCCTGTCCCCCCAGCCACCACACCAAGCTGAAGAAGACATGGCTCACA	540
Db	1090	TCTGGCCCAAGGGCCCTGTCCCCCCAGCCACCACACCAAGCTGAAGAAGACATGGCTCACA	1149
QY	541	CGSCACTCGAGCAGTTTGAATGTCCACGGGCTGCCCTGAGTcGAGGAGAGGCCCGGTT	600
Db	1150	CGSCACTCGAGCAGTTTGAATGTCCACGGGCTGCCCTGAGGCGCCAGGAGAGGCCCGGTT	1209
QY	601	GCTCGGCTCCGGGCCCTCAAAGGGCAGGAGCCCCGAGGTCCAGGGAGCAATGGGCAGT	660
Db	1210	GCTCAGCTCCGGGCCCTCAAAGGGCAGGAGCCCCGAGATCCAAGGAGCAGTGGGCGGC	1269
QY	661	CCAGCCCCAAGCGGCCACCGGACCCCTTTCCAGGCACTGCAGAAACAGGGGGCTGGGGGT	720
Db	1270	CCAGCCCCAAGCGGCCACCGGACCCCTTTTCAGGCACTGCAGAAACAGGGGGCTGGGGGT	1329
QY	721	TGGCAGGAGGTTCCGGGACACATCGATAGGGAAACAAGGATGTGGACTCGGGACAGCATGAT	780
Db	1330	TGGCAGGAGGTACGGGACACATCGATAGGGAAACAAGGAGCGGACTCGGGACAGCATGAT	1389

QY	781	GACCAGAAAGGACCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTTCAGGACATA	840
Db	1390	GACCAGAGAGGACCCCGAGACGGCCAGGCCAGTCTCCAGGACCCAGGACTTTCAGGAAATA	1449
QY	841	CCATGCCTGCTTCTCCCTGCAAAAACCTGGCTCAATGCCAAAGTTGTGCCAGGCAGCTGGA	900
Db	1450	CCATGCCTGGCTCTCCCTGCAAAAACCTGGCTCAATGCCAAAGTTGTGCCAGGCAGCTGGA	1509
QY	901	GAGGAGGAGGGCACGCCTTGCCACTCTCAGCAAGTGCAGATCGCCTCTGGGAGGGAG	960
Db	1510	GAGGAGGAGGGCCCCCGCGCCCACTTTCAGCAAGTGCAGATCACCTCTGGGAGGGGAA	1569
QY	961	CTGCAGCAGGAGGAAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCAGGTCCTGGC	1020
Db	1570	CTGCAGCAGGAGGAAGACACAGCCGCCAACTCCAGCTCTGAGGAAGGCCAGGTCCTGGC	1629
QY	1021	CCTGACAGCCGGCTCAGCACAGGCCTCGCCAAAGCACCTGCTCAGTGGTTTGGGGACCGA	1080
Db	1630	CCTGATGGCAGGCTCAGCACAGGCCTCTCCAAGCACCTGCTCAGTAGTCTAGGGGACCGA	1689
QY	1081	CTGTGCCGCTGCTGCGGAGGGAGCGGGAGGCCCTGGCTTGGCCCCAGCGGAAGGCCAA	1140
Db	1690	CTGTGCCGCTGCTGCGGAGAGAGCGGGAAAGCCCTGGCCCTGGCCCCAGCGGAAGGCCAA	1749
QY	1141	GGCCAGCCCGTGACAGGGGACAGCCAGGCATTCACGCTGCTGACGCCGTTGCCACCAT	1200
Db	1750	GGCCAGCCCGTGACAGAGGACAAACCAGGCATTCACGCTGCTGACGCCGTTGCCACCA	1809
QY	1201	GGACTCTTCAACACCCACTGGCGATGTCCCCTGTCAGCCACGGCTGTGTGTGGCCTGT	1260
Db	1810	GGACTCTTCAACACCCACTGGCGATGTCCCCTGTCAGCCACGGCTGTGTGTGGCCTGC	1869
QY	1261	GGTCTGTGGCAGGCACCTGGCGGGCCAGGGAGAAAGCAGGCTTTCAGGAGCAGTCCCG	1320
Db	1870	GGTCTGTGGCAGGCCTGGCGGGCCAGGGAGAAAGCAGGCTTTCGGGAGCAGTCCACG	1929
QY	1321	GAGGAGTGACGCAGGAGCGCGGCACGCTGCCCTGTTCCTGATGCTGACCCAGTTTGTG	1380
Db	1930	GAGGAGTGACGCAGGAGGCTGGGCAGCTGCCCTGTTCCTGACGCTGACCCAGTTTGTG	1989
QY	1381	TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACAGGTCTGGGTCAAGTTTGAT	1440
Db	1990	TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACAGGTCTGGGTCAAGTTTGAT	2049
QY	1441	ATCCGGGGCACTGCCCTGCCAAGCTGATGCCGGGTATGGCCCCCGGGGATGCAGGC	1500
Db	2050	ATCCGGGGCACTGCCCTGCCAAGCTGATGCCGGGTATGGCCCCCGGGGATGCAGGC	2109
QY	1501	CAGCAGAAGGAATCAACACAGAAACGCCCCCAACTCCACAACCTTCTGCAATGGCGAC	1560
Db	2110	CAGCAGAAGGAATCAACACAGAAACGCCCCCAACTCCACAACCTTCTGCAATGGCGAC	2169
QY	1561	ACCCACAGGACCAAGAGCATCAAAGAGGAGACCCCGATTCCCGTGAGACCCCGAGCAG	1620
Db	2170	ACCCACAGGACCAAGAGCATCAAAGAGGAGACCCCGATTCCACTGAGACCCCGAGCAG	2229
QY	1621	GACCGTGTGGCCGAGGGGCCCTTGTCTTCTCTCTGCGAACTGCTGGCTTCTACC	1680
Db	2230	GACCGTGTGGCCAAAGCGCCCTGCTTGTCTTCTCTCTGCGAACTGCTGGCTTCTACC	2289
QY	1681	GCGGTCAAACCTCTGCTTGGGCCATGACGGAATACACATGGCCCTTCGCCCCCGCTCCTCCG	1740
Db	2290	GCGGTCAAACCTCTGCTTGGGGCATGAGCGGATAACATGGCCCTTCGCCCCCGCTCCTCCG	2349
QY	1741	GCCCTGCCCAGTGATGACCGCATCAACAAACATCCTGGACAGCATATTCGCACAGGTGGTG	1800
Db	2350	GCCCTGCCCAGTGATGACCGCATCAACAAACATCCTGGACAGCATATTCGCACAGGTGGTG	2409
QY	1801	GAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGGCTTCGAGCTGGCCCCGGGTCTGCGC	1860
Db	2410	GAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGGCTTCGAGCTGGCCCCGGGTCTGCGC	2469
QY	1861	AAGGGCCTGGGCGCTGCCCTCTCTCCAGTGGGGCCCCGGCTGCCCTCCCCCAGGGGCTTG	1920

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 11:46:34 ; Search time 1179 Seconds
(without alignments)
11537.515 Million cell updates/sec

Title: US-10-024-368-1
Perfect score: 3202
Sequence: 1 ttttactacaaggatccgag.....tgtgccccttttatgtacagg 3202

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3202	100.0	3202	6	AAD31072 Alternati
2	3192.4	99.7	3202	6	AAD31077 Human Hai
3	3190.8	99.7	3202	8	ACD26351 Human Hai
4	3185	99.5	5709	7	ABX13987 Human hai
5	2937.6	91.7	3567	2	AAZ06349 Human Hai
6	533.2	16.7	538	7	ACA57497 Human adi
7	86.8	2.7	335	6	ABS69028 Novel mur
8	66	2.1	429	7	ABX41049 Bovine ES
9	62.4	1.9	114955	2	AAX53491 Human ade
10	60	1.9	2000	7	ADA71938 Rice gene
11	56.2	1.8	2000	7	ADA71938 Rice gene
12	55.4	1.7	65	6	ABN52357 Mouse spl
13	54.6	1.7	4783	3	AAC76742 Human ORF
14	52	1.6	1402	2	AAV36085 DNA encod
15	50	1.6	10732	3	AAA10594 Gene enco
16	49	1.5	125401	4	AAD17186 Streptomy
17	48.2	1.5	427	5	AAF67464 Novel hum
18	48.2	1.5	479	8	ACH34443 Human end
19	48.2	1.5	2359	4	AAI61114 Human pol
20	48.2	1.5	3044	6	ABQ54675 Human ova
21	48.2	1.5	4628	4	AAI59328 Human pol
22	48.2	1.5	6204	5	ABV27441 Human pro
23	48.2	1.5	6204	5	ABV21621 Human pro

24	48.2	1.5	6995	10	ADE40508	Ade40508 Human nuc
25	48.2	1.5	110000	4	AAI99682_39	Continuation (40 o
26	48	1.5	114955	2	AAX53491	Aax53491 Human ade
27	47.8	1.5	1089	2	AAZ32065	Aaz32065 Human MET
28	47.8	1.5	1089	5	AAC90322	Aac90322 I08144 CD
29	47.6	1.5	6741	3	AAA10595	Aaa10595 Gene enco
30	47.4	1.5	339	2	AAQ03321	Aaq03321 Eimeria t
31	47.4	1.5	747	2	AAQ03323	Aaq03323 Eimeria t
32	47.4	1.5	954	1	AAN90505	Aan90505 DNA encod
33	47.4	1.5	957	2	AAV66836	Aav66836 Group B E
34	47.4	1.5	957	2	AAZ32063	Aaz32063 Human MET
35	47.4	1.5	957	5	AAC90320	Aac90320 X15898 CD
36	47.4	1.5	1089	1	AAN92576	Aan92576 Sequence
37	47.4	1.5	1089	2	AAT93593	Aat93593 Eimeria t
38	47.2	1.5	2386	2	AAQ63477	Aaq63477 Micrococc
39	46.2	1.4	1300	6	ABK83980	Abk83980 Human CDN
40	46.2	1.4	4309	6	ABV77882	Abv77882 Hypoxia-i
41	46.2	1.4	4309	9	ADD18798	Add18798 Human dis
42	46.2	1.4	4697	4	AAK51593	Aak51593 Human pol
43	46.2	1.4	4746	4	AAK52577	Aak52577 Human pol
44	46.2	1.4	10678	4	AAI05144	Aai05144 Human rep
45	46.2	1.4	10678	4	ABL98032	Abl98032 Human tes

ALIGNMENTS

RESULT 1
AAD31072
ID AAD31072 standard; DNA; 3202 BP.
XX
AC AAD31072;
XX
DT 18-JUN-2002 (first entry)
XX
DE Alternative version of human Hairless gene (HR).
XX
KW Human; Hairless protein; Hr protein; HR gene; dermatological condition;
KW hair loss; gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key
CDS
FT
FT /*tag= a
FT /product= "Human Hr protein #1"
FT /transl_except= (pos:850..852, aa:Ala)
FT /note= "CDS does not include start codon"
FT /partial
XX
US6348348-B1.
XX
PD 19-FEB-2002.
XX
PF 07-APR-1999; 99US-00287354.
XX
PR 07-APR-1998; 98US-0080888P.
XX
PA (CARN-) CARNEGIE INST WASHINGTON.
XX
PI Thompson CC;
XX
XX WPI; 2002-204622/26.
DR P-PSDB; AAE19794.
XX
PT Novel expression construct, useful in the diagnosis and treatment of
PT dermatological conditions, such as hair loss, contains a Hairless gene
PT sequence.
XX
PS Claim 4; Col 35-38; 48pp; English.
XX
CC The invention relates to human Hairless (Hr) polypeptides and nucleic
CC acid molecules (HR) encoding such polypeptides. The invention also

CC relates to the Hairless expression constructs which may be used in
CC transcription assays. Sequences of the invention are used in diagnosis
CC and treatment of dermatological conditions such as hair loss. They are
CC also used in gene therapy. Polynucleotides of the invention can be used
CC as probes for the detection of hair loss. The present sequence is an
CC alternative version of human Hairless gene (HR). Note: This sequence is
CC stated as being the same as that shown as SEQ ID NO:1 in Column 25-32
CC (AAD31077) of the specification. However the sequences differ at
CC positions.
XX
SQ Sequence 3202 BP; 599 A; 1040 C; 1029 G; 534 T; 0 U; 0 Other;

Query Match 100.0%; Score 3202; DB 6; Length 3202;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTACTACAGGATCCGAGCATTCCTCCAGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA 60
Db 1 TTTTACTACAGGATCCGAGCATTCCTCCAGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA 60

QY 61 CCTGGGTTGTTGGCTTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGAACGC 120
Db 61 CCTGGGTTGTTGGCTTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGAACGC 120

QY 121 CCTTCACTGCACAGAGGGATGGAGATGGAGCTGGCCAGCGATGGGAACCTCTTGCCCG 180
Db 121 CCTTCACTGCACAGAGGGATGGAGATGGAGCTGGCCAGCGATGGGAACCTCTTGCCCG 180

QY 181 CTCTTCTGGGAGCCAGACACTGTGCCCTGGACCTCTGTGCCCGCTGTCCCCAGGC 240
Db 181 CTCTTCTGGGAGCCAGACACTGTGCCCTGGACCTCTGTGCCCGCTGTCCCCAGGC 240

QY 241 CTGTGTTTATCTCTTGGCAACGTCTGGGCTGGCCAGGCGATGGGAACCTTGGGTACCAG 300
Db 241 CTGTGTTTATCTCTTGGCAACGTCTGGGCTGGCCAGGCGATGGGAACCTTGGGTACCAG 300

QY 301 CTGGGGCCACCAACACCAAGGTGCCCTCTCCTGAGCCCGCTGTCAACCCAGCGGGC 360
Db 301 CTGGGGCCACCAACACCAAGGTGCCCTCTCCTGAGCCCGCTGTCAACCCAGCGGGC 360

QY 361 TGCTGTTTATCTTACCCACCCACTAAAGTTGGGATCTTGGCCCTTGTGGGAAGTGCCAG 420
Db 361 TGCTGTTTATCTTACCCACCCACTAAAGTTGGGATCTTGGCCCTTGTGGGAAGTGCCAG 420

QY 421 GAGGGCCTGAGGGGGTGCCAGTGGAGCCAGCAACCCAGCGAGGAAGTGAACAAGGCC 480
Db 421 GAGGGCCTGAGGGGGTGCCAGTGGAGCCAGCAACCCAGCGAGGAAGTGAACAAGGCC 480

QY 481 TCTGGCCCCAGGGCCTGTCCCCCAGCCACCAACCAAGCTGAAGAAGCATGGCTACA 540
Db 481 TCTGGCCCCAGGGCCTGTCCCCCAGCCACCAACCAAGCTGAAGAAGCATGGCTACA 540

QY 541 CGGCACCTCGGAGCAGTTTGAATGTCCACGGCGTGCCCTGAGGTCGAGGAGAGGCCGGTT 600
Db 541 CGGCACCTCGGAGCAGTTTGAATGTCCACGGCGTGCCCTGAGGTCGAGGAGAGGCCGGTT 600

QY 601 GCTCGGCTCCGGGCCCTCAAAGGGCAGGACGCCCGGAGGTCCAGGGAGCAATGGGCAGT 660
Db 601 GCTCGGCTCCGGGCCCTCAAAGGGCAGGACGCCCGGAGGTCCAGGGAGCAATGGGCAGT 660

QY 661 CCAGCCCCAAGCGGCCACCGGACCTTTCCAGGCACTGCAGAACAGGGGCTGGGGT 720
Db 661 CCAGCCCCAAGCGGCCACCGGACCTTTCCAGGCACTGCAGAACAGGGGCTGGGGT 720

QY 721 TGGCAGGAGGTTCCGGACACATCGATAGGGAACAAGGATGTGGACTCGGACAGCATGAT 780
Db 721 TGGCAGGAGGTTCCGGACACATCGATAGGGAACAAGGATGTGGACTCGGACAGCATGAT 780

QY 781 GAGCAGAAAGGACCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 840
Db 781 GAGCAGAAAGGACCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 840

QY 841 CCATGCCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCCCAGGCAGCTGGA 900

Db 841 CCATGCCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCAGGCAGCTGGA 900

QY 901 GAGGAGGAGGGCACGCCTGCCACTCTCAGCAAGTGGGAGATCGCCTCTGGGAGGGGAG 960
Db 901 GAGGAGGAGGGCACGCCTGCCACTCTCAGCAAGTGGGAGATCGCCTCTGGGAGGGGAG 960

QY 961 CTGCAGCAGGAGGAAGACACAGCCCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1020
Db 961 CTGCAGCAGGAGGAAGACACAGCCCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1020

QY 1021 CCTGACAGCCGGCTCAGCACAGGCCTCGCCAAGCACCTGCTCAGTGTGGGGGACCGA 1080
Db 1021 CCTGACAGCCGGCTCAGCACAGGCCTCGCCAAGCACCTGCTCAGTGTGGGGGACCGA 1080

QY 1081 CTGTGCCGCTGCTGCGAGGGAGCGGAGGCCCTGGCTTGGGCCAGCGGGAAGGCCAA 1140
Db 1081 CTGTGCCGCTGCTGCGAGGGAGCGGAGGCCCTGGCTTGGGCCAGCGGGAAGGCCAA 1140

QY 1141 GGGCCAGCCGTGACAGGGGACAGCCAGGCATTCACGCTGCTGCAGCGTTGCCACCAT 1200
Db 1141 GGGCCAGCCGTGACAGGGGACAGCCAGGCATTCACGCTGCTGCAGCGTTGCCACCAT 1200

QY 1201 GGACTCTTCAACACCCACTGGCGATGTCCCGCTGCAGCCACCGGCTGTGTGGCCTGT 1260
Db 1201 GGACTCTTCAACACCCACTGGCGATGTCCCGCTGCAGCCACCGGCTGTGTGGCCTGT 1260

QY 1261 GGTCTGTGGCAGGCACCTGGCGGGCCAGGGAGAAAGCAGGCTTTCAGGAGCAGTCCGG 1320
Db 1261 GGTCTGTGGCAGGCACCTGGCGGGCCAGGGAGAAAGCAGGCTTTCAGGAGCAGTCCGG 1320

QY 1321 GAGGAGTGACCGCAGGAGGCCGGGCACGCTGCCTGTTCCTGATGCTGACCCAGTTTGT 1380
Db 1321 GAGGAGTGACCGCAGGAGGCCGGGCACGCTGCCTGTTCCTGATGCTGACCCAGTTTGT 1380

QY 1381 TCCAGCCAGGCTTGGCAGAGCTGAGCACTGCAATGCACCAAGTCTGGGTCAAGTTTGT 1440
Db 1381 TCCAGCCAGGCTTGGCAGAGCTGAGCACTGCAATGCACCAAGTCTGGGTCAAGTTTGT 1440

QY 1441 ATCCGGGGCACTGCCCCCTGCCAAGCTGATGCCCGGGTATGGGCCCGGGGATGAGGC 1500
Db 1441 ATCCGGGGCACTGCCCCCTGCCAAGCTGATGCCCGGGTATGGGCCCGGGGATGAGGC 1500

QY 1501 CAGCAGAAGGAATCAACACAGAAAACGCCCCCAACTCCACAACCTTCTGCAATGGCGAC 1560
Db 1501 CAGCAGAAGGAATCAACACAGAAAACGCCCCCAACTCCACAACCTTCTGCAATGGCGAC 1560

QY 1561 ACCCAGAGCAACAGAGCATCAAAGAGAGACCCCCCGATTCCGCTGAGACCCAGCAGAG 1620
Db 1561 ACCCAGAGCAACAGAGCATCAAAGAGAGACCCCCCGATTCCGCTGAGACCCAGCAGAG 1620

QY 1621 GACCGTGTGCGGAGGGCCCTGCTTGTTCCTTCTCTGCGAATGCTGGCTTCTACC 1680
Db 1621 GACCGTGTGCGGAGGGCCCTGCTTGTTCCTTCTCTGCGAATGCTGGCTTCTACC 1680

QY 1681 GCGGTCAAACCTCTGCTTGGGCCATGAGCGAATACACATGGCCTTCGCCCCCGTCACTCCG 1740
Db 1681 GCGGTCAAACCTCTGCTTGGGCCATGAGCGAATACACATGGCCTTCGCCCCCGTCACTCCG 1740

QY 1741 GCCCTGCCAGTGATGACCGCATCACCAACATCTTGACAGCATTTATCGCACAGGTGGTG 1800
Db 1741 GCCCTGCCAGTGATGACCGCATCACCAACATCTTGACAGCATTTATCGCACAGGTGGTG 1800

QY 1801 GAACGGAAGATCCAGGAGAAAGCCCTGGGCGCGGGCTTCGAGCTGGCCCGGGTCTGCGC 1860
Db 1801 GAACGGAAGATCCAGGAGAAAGCCCTGGGCGCGGGCTTCGAGCTGGCCCGGGTCTGCGC 1860

QY 1861 AAGGGCCTGGGCTGCCCTCTCTCCAGTGGGGCCCCGGCTGCCTCCCCCAGGGCTTTG 1920
Db 1861 AAGGGCCTGGGCTGCCCTCTCTCCAGTGGGGCCCCGGCTGCCTCCCCCAGGGCTTTG 1920

QY 1921 CTGTGGCTGCAGGAGCCCCAGCCTTGGCCTCGGGCTGGCTTCCACCTCTTCAGGAGCAC 1980

Db 506 EGGHACHSQVRRSPGLGGLQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 565
QY 361 LCRLRREREALAWAQREGQPAVTGDSPIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 566 LCRLRGEREALAWAQREGQPAVTEDSPIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 625
QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEELSTAMHVWVKFD 480
Db 626 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEELSTAMHVWVKFD 685
QY 481 IRGHCPQADARVWAPGDAGQOKESTQKTPPTQPSCNGDTHRTKSIKEETPDSAEAPAE 540
Db 686 IRGHCPQADARVWAPGDAGQOKESTQKTPPTQPSCNGDTHRTKSIKEETPDSAEAPAE 745
QY 541 DRAGRGPLPCPSLCELLASTAVKLCGHERIHMAFAPVTPALPSDDRITNILDIIAQVV 600
Db 746 DRAGRGPLPCPSLCELLASTAVKLCGHDRIHMAFAPVTPALPSDDRITNILDIIAQVV 805
QY 601 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQPCPRRGFHLFQEH 660
Db 806 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQPCPRRGFHLFQEH 865
QY 661 WRQGPVLVSGIORTLQGNLWGTALGALGGVQVQALSPLGPPQPSLGSSTTFWEGFSWPE 720
Db 866 WRQGPVLVSGIORTLQGNLWGTALGALGGVQVQALSPLAPPQPSLGSSTTFWEGFSWPE 925
QY 721 LRPKSDGSGVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 780
Db 926 LRPKSDGSGVLLHRLAFGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 985
QY 781 POLWAAAYGVSPHRGHLGTKNLCVEVADLVSIHVADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 986 POLWAAAYGVSPHRGHLGTKNLCVEVADLVSIHVARTPLPAWHRAQKDFLSGLDGEGLWS 1045
QY 841 PGSQVSTVWHVFRQAQDAQRIIRFLQWVCPAGAGALEPGAPGSCVLDAGLRRRLREEWGVS 900
Db 1046 PGSQVSTVWHVFRQAQDAQRIIRFLQWVCPAGAGALEPGAPGSCVLDAGLRRRLREEWGVS 1105
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSACLCHQGPSLPPDCHL 960
Db 1106 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSACLCHQGPSLPPDCHL 1165
QY 961 LYAQMDWAVFQAVKVAVGTLQEA 984
Db 1166 LYAQMDWAVFQAVKVAVGTLQEA 1189

RESULT 5

US-10-024-368-6
; Sequence 6, Application US/10024368
; Publication No. US20030027300A1

GENERAL INFORMATION:

; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mouse
US-10-024-368-6

Query Match 76.6%; Score 4102.5; DB 14; Length 1182;
Best Local Similarity 78.4%; Pred. No. 1.8e-282;
Matches 771; Conservative 62; Mismatches 144; Indels 7; Gaps 5;

QY 1 FYYKDPISIPRLAKEPLAAAPGLFGLNSGGHLQRAHFAERPSLHQRDGEMGAGRQONPCP 60
Db 206 FYHKDENILRPAKEPL--AESGMLGAPGGHLQACSESGPSLHQRDGETGAGRQONLCP 263
QY 61 LFLGQPDTPVWTWPACPPGLVHTILGNVWAGPDGNLYQLGPPATPRCPSPPEPVTQRG 120
Db 264 VFLGYPDTPVPRAPWPSCPPGLVHSLGNIWAGPGNSLGYQLGPPATPRCPSPGPPTFPFG 323
QY 121 CCSSYPPTKGGDLGPCCKQEGLEGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180
Db 324 CCSSHLPAREGDLGPCRKQDQSPGSGSGPSESSEERNKADS-RACPPSHHTKLKKTWLT 382
QY 181 RHSEQFECPRGCEVEERPVARLRALKRAGSPEVQGAMGSPAPKRPPDPFPPTAEQAGG 240
Db 383 RHSEQFECPGGCGKEESSATGLRALKRAGSPEVQGASRGAPKRPSHPFPGTGRQGARA 442
QY 241 WQEVRTDSIGNKDVDSGQHDQKGPQDQASLQDPGLQDIPCLALPAKLAQCQSCQAAG 300
Db 443 WQETPETIIGSK-AEAEQOEEOGRGPRDGRIRLQESRLVDTSCQHLAGVTQCQSCVQAAG 501
QY 301 EGGHACHSQVRRSPGLGGLQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360
Db 502 EVGVLTHSQSKRRSPL-BEKQLEEDSSATSEEGGGGPGPEASLNKGLAKHLLSGLGDR 560
QY 361 LCRLRREREALAWAQREGQPAVTGDSPIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 561 LCRLRREREALAWAQREGQPAVTGDSPIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 620
QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEELSTAMHVWVKFD 480
Db 621 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEELSTAMHVWVKFD 680
QY 481 IRGHCPQADARVWAPGDAGQOKESTQKTPPTQPSCNGDTHRTKSIKEETPDSAEAPAE 540
Db 681 IRGHCPQADARVWAPGDAGQOKESTQKTPPTQPSCNGDTHRTKSIKEETPDSAEAPAE 740
QY 541 DRAGRGPLPCPSLCELLASTAVKLCGHERIHMAFAPVTPALPSDDRITNILDIIAQVV 600
Db 741 DRAGRGPLPCPSLCELLASTAVKLCGHDRIHMAFAPVTPALPSDDRITNILDIIAQVV 800
QY 601 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQPCPRRGFHLFQEH 660
Db 801 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQPCPRRGFHLFQEH 858
QY 661 WRQGPVLVSGIORTLQGNLWGTALGALGGVQVQALSPLGPPQPSLGSSTTFWEGFSWPE 720
Db 859 WRQGPVLVSGIORTLQGNLWGTALGALGGVQVQALSPLGPPQPSLGSSTTFWEGFSWPE 918
QY 721 LRPKSDGSGVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 780
Db 919 TRPKLDEGSGVLLHRLTGLDKDASRVQNLASSLPLPEYCAHQKLNLASYLPLGLTLHPLE 978
QY 781 POLWAAAYGVSPHRGHLGTKNLCVEVADLVSIHVADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 979 POLWAAAYGVSPHRGHLGTKNLCVEVADLVSIHVADTPLPAWHRAQKDFLSGLDGEGLWS 1038
QY 841 PGSQVSTVWHVFRQAQDAQRIIRFLQWVCPAGAGALEPGAPGSCVLDAGLRRRLREEWGVS 900
Db 1039 PGSQVSTVWHVFRQAQDAQRIIRFLQWVCPAGAGALEPGAPGSCVLDAGLRRRLREEWGVS 1098
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSACLCHQGPSLPPDCHL 960
Db 1099 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSACLCHQGPSLPPDCHL 1158
QY 961 LYAQMDWAVFQAVKVAVGTLQEA 984
Db 1159 LYAQMDWAVFQAVKVAVGTLQEA 1182

RESULT 6

US-10-024-368-5
; Sequence 5, Application US/10024368
; Publication No. US20030027300A1

QY 961 LYAQMDWAVFQAVKVAVGTLOEAK 984
Db 1166 LYAQMDWAVFQAVKVAVGTLOEAK 1189

RESULT 3

US-10-122-013-17
; Sequence 17, Application US/10122013
; Publication No. US20030077614A1
; GENERAL INFORMATION:
; APPLICANT: Cristiano, Angela
; TITLE OF INVENTION: NUCLEIC ACIDS FOR INHIBITING HAIRLESS PROTEIN EXPRESSION AND METH
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: 0575/62637A
; CURRENT APPLICATION NUMBER: US/10/122,013
; CURRENT FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-122-013-17

Query Match 99.5%; Score 5331; DB 14; Length 1189;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 981; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FYYKDPSSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERPSLHQRDGEMGAGRQONPCP 60
Db 206 FYYKDPSSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERPSLHQRDGEMGAGRQONPCP 265
QY 61 LFLGQPDTPWTSWPACPPGLVHTLGNVWAGPGDGNLGYQLGPPATPRCPSPEPPVTORG 120
Db 266 LFLGQPDTPWTSWPACPPGLVHTLGNVWAGPGDGNLGYQLGPPATPRCPSPEPPVTORG 325
QY 121 CCSSYPPTKGGDLGPCGKCEGLEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180
Db 326 CCSSYPPTKGGDLGPCGKCEGLEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 385
QY 181 RHSEQFECPRGCPEVEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFGTAEOGAGG 240
Db 386 RHSEQFECPRGCPEVEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFGTAEOGAGG 445
QY 241 WQEVRTDSIGNKDVDSGQHQDEQKGPQDGOASLQDPGLQDIPCLALPAKLAQCQSCAQAG 300
Db 446 LQEVRTDSIGNKDVDSGQHQDEQKGPQDGOASLQDPGLQDIPCLALPAKLAQCQSCAQAG 505
QY 301 EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSLGLDR 360
Db 506 EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSLGLDR 565
QY 361 LCRLRREREALAWAQREGQGPVTDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 566 LCRLRREREALAWAQREGQGPVTDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 625
QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEI STAMHQVWKFD 480
Db 626 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEI STAMHQVWKFD 685
QY 481 IRGHCPCQADARVWAPGDAGQOKESTQKTPPTQPSGNGDTHRTKSIKEETPD SAETPAE 540
Db 686 IRGHCPCQADARVWAPGDAGQOKESTQKTPPTQPSGNGDTHRTKSIKEETPD SAETPAE 745
QY 541 DRAGRGLPCPSLCELLASTAVKLCGHERIHMAFAPVTPALPDDRITNILD SI AQVV 600
Db 746 DRAGRGLPCPSLCELLASTAVKLCGHERIHMAFAPVTPALPDDRITNILD SI AQVV 805
QY 601 ERKIQEKALGPGLRAGPLRKGLGLPLSPVRPRLPPPGALLWLQEPQPCPRRGHFLFQEH 660
Db 806 ERKIQEKALGPGLRAGPLRKGLGLPLSPVRPRLPPPGALLWLQEPQPCPRRGHFLFQEH 865

QY 661 WRQGPVLVSGIORTLQGNLWGTEALGALGGQVQALSPLPPQPSLSGSTTFWEGFSWPE 720
Db 866 WRQGPVLVSGIORTLQGNLWGTEALGALGGQVQALSPLPPQPSLSGSTTFWEGFSWPE 925
QY 721 LRPKSDGSVLLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 780
Db 926 LRPKSDGSVLLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 985
QY 781 PQLWAAAYGVSPHRGHLGTKNLCVEVADLVSVLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 986 PQLWAAAYGVSPHRGHLGTKNLCVEVADLVSVLVHADTPLPAWHRAQKDFLSGLDGEGLWS 1045
QY 841 PGSQVSTVWHVFRADOAQRIRRFLOMVCAPAGAGALEPGAPGSCVLDAGLRRRLREEWGVS 900
Db 1046 PGSQVSTVWHVFRADOAQRIRRFLOMVCAPAGAGALEPGAPGSCVLDAGLRRRLREEWGVS 1105
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSACLCHQGPSLPPDCHL 960
Db 1106 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSACLCHQGPSLPPDCHL 1165
QY 961 LYAQMDWAVFQAVKVAVGTLOEAK 984
Db 1166 LYAQMDWAVFQAVKVAVGTLOEAK 1189

RESULT 4

US-10-024-368-3
; Sequence 3, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-3

Query Match 98.7%; Score 5290; DB 14; Length 1189;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 974; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 FYYKDPSSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERPSLHQRDGEMGAGRQONPCP 60
Db 206 FYYKDPSSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERPSLHQRDGEMGAGRQONPCP 265
QY 61 LFLGQPDTPWTSWPACPPGLVHTLGNVWAGPGDGNLGYQLGPPATPRCPSPEPPVTORG 120
Db 266 LFLGQPDTPWTSWPACPPGLVHTLGNVWAGPGDGNLGYQLGPPATPRCPSPEPPVTORG 325
QY 121 CCSSYPPTKGGDLGPCGKCEGLEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180
Db 326 CCSSYPPTKGGDLGPCGKCEGLEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 385
QY 181 RHSEQFECPRGCPEVEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFGTAEOGAGG 240
Db 386 RHSEQFECPRGCPEVEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFGTAEOGAGG 445
QY 241 WQEVRTDSIGNKDVDSGQHQDEQKGPQDGOASLQDPGLQDIPCLALPAKLAQCQSCAQAG 300
Db 446 LQEVRTDSIGNKDVDSGQHQDEQKGPQDGOASLQDPGLQDIPCLALPAKLAQCQSCAQAG 505
QY 301 EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSLGLDR 360

QY 181 RHSEQFECPRGCEVEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFFGTAEQAGG 240
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-4

181 RHSEQFECPRGCEVEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFFGTAEQAGG 240

QY 241 WQEVDRDTSIGNKDVDSGQHDQKGPQDQASLQDPLCLALPAKLAQCSCAAAG 300

241 WQEVDRDTSIGNKDVDSGQHDQKGPQDQASLQDPLCLALPAKLAQCSCAAAG 300

QY 301 EGGGHACHSQVRRSPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360

301 EGGGHACHSQVRRSPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360

QY 361 LCRLLRREREALAWAQREGQGPVATGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420

361 LCRLLRREREALAWAQREGQGPVATGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420

QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 480

421 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 480

QY 481 IRGHCPQADARVWAPGDAGQKKESTQKTPTPOPSCNGDTHRTKSIKEETPDSAETPAE 540

481 IRGHCPQADARVWAPGDAGQKKESTQKTPTPOPSCNGDTHRTKSIKEETPDSAETPAE 540

QY 541 DRAGRGLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDDRITNILDIIAQVV 600

541 DRAGRGLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDDRITNILDIIAQVV 600

QY 601 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQCPRRGFHLFQEH 660

601 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQCPRRGFHLFQEH 660

QY 661 WRQGPVLVSGIORTLQGNLWGTEALGAGGVQVQALSPLGPPQSSSLGSTTFWEGFSWPE 720

661 WRQGPVLVSGIORTLQGNLWGTEALGAGGVQVQALSPLGPPQSSSLGSTTFWEGFSWPE 720

QY 721 LRPKSDGGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 780

721 LRPKSDGGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 780

QY 781 POLWAAAYGVSPHRGHLGKTNLCVEVADLVSVILVHADTPLPAWHRAQKDFLSGLDGEGLWS 840

781 POLWAAAYGVSPHRGHLGKTNLCVEVADLVSVILVHADTPLPAWHRAQKDFLSGLDGEGLWS 840

QY 841 PGSQVSTVWHVFRQAQDAQIRIRFLQMVCPAGAGALEPGAGSCVLDAGLRRRLREEWGVS 900

841 PGSQVSTVWHVFRQAQDAQIRIRFLQMVCPAGAGALEPGAGSCVLDAGLRRRLREEWGVS 900

QY 901 CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSTQHFLSPETSALSACLCHQGPSLPPDCHL 960

901 CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSTQHFLSPETSALSACLCHQGPSLPPDCHL 960

QY 961 LYAQMDWAVFQAVKVAVGTLQEAK 984

961 LYAQMDWAVFQAVKVAVGTLQEAK 984

RESULT 2

US-10-024-368-4
; Sequence 4, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-4

Query Match 99.7%; Score 5343; DB 14; Length 1189;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 982; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FYKDPSPRLAKEPLAAEPGLFGLNSGGHLQORAGEERPSSLHQRDGEMGAGRQONPCP 60

Db 206 FYKDPSPRLAKEPLAAEPGLFGLNSGGHLQORAGEERPSSLHQRDGEMGAGRQONPCP 265

QY 61 LFLGQPDTPVWTWPACPPGLVHTLGNVWAGDGNLGYQLGPPATPRCPSPEPPVTQRG 120

Db 266 LFLGQPDTPVWTWPACPPGLVHTLGNVWAGDGNLGYQLGPPATPRCPSPEPPVTQRG 325

QY 121 CCSSYPPTKGGDLGPCGKCEGLEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180

Db 326 CCSSYPPTKGGDLGPCGKCEGLEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 385

QY 181 RHSEQFECPRGCEVEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFFGTAEQAGG 240

Db 386 RHSEQFECPRGCEVEERPVARLRALKRAGSPEVQGMGSPAPKRPPDPFFGTAEQAGG 445

QY 241 WQEVDRDTSIGNKDVDSGQHDQKGPQDQASLQDPLCLALPAKLAQCSCAAAG 300

Db 446 WQEVDRDTSIGNKDVDSGQHDQKGPQDQASLQDPLCLALPAKLAQCSCAAAG 505

QY 301 EGGGHACHSQVRRSPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360

Db 506 EGGGHACHSQVRRSPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 565

QY 361 LCRLLRREREALAWAQREGQGPVATGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420

Db 566 LCRLLRREREALAWAQREGQGPVATGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 625

QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 480

Db 626 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 685

QY 481 IRGHCPQADARVWAPGDAGQKKESTQKTPTPOPSCNGDTHRTKSIKEETPDSAETPAE 540

Db 686 IRGHCPQADARVWAPGDAGQKKESTQKTPTPOPSCNGDTHRTKSIKEETPDSAETPAE 745

QY 541 DRAGRGLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDDRITNILDIIAQVV 600

Db 746 DRAGRGLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDDRITNILDIIAQVV 805

QY 601 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQCPRRGFHLFQEH 660

Db 806 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQCPRRGFHLFQEH 865

QY 661 WRQGPVLVSGIORTLQGNLWGTEALGAGGVQVQALSPLGPPQSSSLGSTTFWEGFSWPE 720

Db 866 WRQGPVLVSGIORTLQGNLWGTEALGAGGVQVQALSPLGPPQSSSLGSTTFWEGFSWPE 925

QY 721 LRPKSDGGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 780

Db 926 LRPKSDGGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 985

QY 781 POLWAAAYGVSPHRGHLGKTNLCVEVADLVSVILVHADTPLPAWHRAQKDFLSGLDGEGLWS 840

Db 986 POLWAAAYGVSPHRGHLGKTNLCVEVADLVSVILVHADTPLPAWHRAQKDFLSGLDGEGLWS 1045

QY 841 PGSQVSTVWHVFRQAQDAQIRIRFLQMVCPAGAGALEPGAGSCVLDAGLRRRLREEWGVS 900

Db 1046 PGSQVSTVWHVFRQAQDAQIRIRFLQMVCPAGAGALEPGAGSCVLDAGLRRRLREEWGVS 1105

QY 901 CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSTQHFLSPETSALSACLCHQGPSLPPDCHL 960

Db 1106 CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSTQHFLSPETSALSACLCHQGPSLPPDCHL 1165

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 10:28:55 ; Search time 54 Seconds
(without alignments)
5151.519 Million cell updates/sec

Title: US-10-024-368-2
Perfect score: 5358
Sequence: 1 FYYKDPSPRLAKEPLAAAE.....MDWAVEQAVKVAVGTLQEA 984

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 1166195

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5358	100.0	984	14	US-10-024-368-2
2	5343	99.7	1189	14	US-10-024-368-4
3	5331	99.5	1189	14	US-10-122-013-17
4	5290	98.7	1189	14	US-10-024-368-3
5	4102.5	76.6	1182	14	US-10-024-368-6
6	4067	75.9	1207	14	US-10-024-368-5
7	3084.5	57.6	693	16	US-10-452-858C-93
8	778	14.5	1265	16	US-10-408-765A-795
9	777.5	14.5	2055	12	US-10-193-874-20
10	777.5	14.5	2321	12	US-10-193-874-17
11	777.5	14.5	2358	12	US-10-193-874-14
12	777.5	14.5	2540	12	US-10-193-874-16
13	777.5	14.5	2552	12	US-10-193-874-15
14	711.5	13.3	1417	16	US-10-408-765A-1992
15	696.5	13.0	1212	12	US-10-170-385-19

16	574	10.7	671	15	US-10-264-049-2730	Sequence 2730, Ap
17	564.5	10.5	953	12	US-10-193-874-18	Sequence 18, Appl
18	287.5	5.4	1057	15	US-10-108-260A-3230	Sequence 3230, Ap
19	286.5	5.3	941	12	US-10-424-599-210074	Sequence 210074,
20	284	5.3	747	12	US-10-425-114-38701	Sequence 38701, A
21	278.5	5.2	837	15	US-10-094-749-2698	Sequence 2698, Ap
22	254.5	4.7	579	12	US-10-425-114-56872	Sequence 56872, A
23	249	4.6	363	12	US-10-425-114-42461	Sequence 42461, A
24	247.5	4.6	541	12	US-10-425-114-72403	Sequence 72403, A
25	247	4.6	357	12	US-10-425-114-69939	Sequence 69939, A
26	223	4.2	349	12	US-10-425-114-37462	Sequence 37462, A
27	223	4.2	380	12	US-10-425-114-71856	Sequence 71856, A
28	205	3.8	1561	16	US-10-408-765A-235	Sequence 235, Appl
29	203.5	3.8	19652	15	US-10-084-846A-7	Sequence 7, Appli
30	203.5	3.8	19695	15	US-10-084-846A-3	Sequence 3, Appli
31	203	3.8	1386	14	US-10-327-414-2	Sequence 2, Appli
32	200.5	3.7	19608	15	US-10-084-846A-8	Sequence 8, Appli
33	196	3.7	1564	15	US-10-292-798-1254	Sequence 1254, Ap
34	194	3.6	1259	14	US-10-260-715-8	Sequence 8, Appli
35	189	3.5	1497	12	US-10-016-248-8	Sequence 8, Appli
36	186.5	3.5	19725	15	US-10-084-846A-4	Sequence 4, Appli
37	184	3.4	19723	15	US-10-084-846A-5	Sequence 5, Appli
38	183	3.4	306	12	US-10-425-114-68632	Sequence 68632, A
39	183	3.4	332	12	US-10-425-114-53730	Sequence 53730, A
40	183	3.4	1626	14	US-10-202-167-2	Sequence 2, Appli
41	182.5	3.4	217	12	US-10-424-599-172208	Sequence 172208,
42	182.5	3.4	1684	12	US-10-276-774-2398	Sequence 2398, Ap
43	182	3.4	1430	16	US-10-408-765A-1083	Sequence 1083, Ap
44	181.5	3.4	1857	16	US-10-408-765A-2173	Sequence 2173, Ap
45	180.5	3.4	2161	16	US-10-408-765A-1283	Sequence 1283, Ap

ALIGNMENTS

RESULT 1
US-10-024-368-2
; Sequence 2, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-2

Query Match	100.0%;	Score 5358;	DB 14;	Length 984;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches	984;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	FYYKDPSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAEERPSLHQRDGEMGAGRQONPCP	60	
Db	1	FYYKDPSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAEERPSLHQRDGEMGAGRQONPCP	60	
Qy	61	LFLGQPDTPVWTWPACPPGLVHTLGNVWAGPDGNLGYQLGPPATPRCPSPPEPVTQRG	120	
Db	61	LFLGQPDTPVWTWPACPPGLVHTLGNVWAGPDGNLGYQLGPPATPRCPSPPEPVTQRG	120	
Qy	121	CCSSYPPTKGGDLGPCGKQEGLEGASGASEPSEEVNKAAGPRACPPSHHTKLKKTWLT	180	
Db	121	CCSSYPPTKGGDLGPCGKQEGLEGASGASEPSEEVNKAAGPRACPPSHHTKLKKTWLT	180	

QY 601 ERKIQEKALGPGLRAGPGLRKGLGLPLSPVRPRLLPPPGALLWLQEPQPCPRRGHFLFOEH 660
Db 801 ERKIQEKALGPGLRAGSGLRKGLSLPLSPVTRLSPPGALLWLQEP--PKHGFHLFOEH 858
QY 661 WRQGPVLVSGIQRTLQNLWGTEALGALGGVQVQALSPLGPPQSSLSGTTFWEGFSWPE 720
Db 859 WRQGPVLVSGIQKTLRLSLWGMEALGTLGGVQVQTLTALGPPQPTNLDSTAFWEGFSHPE 918
QY 721 LRPKSDGVSLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 780
Db 919 TRPKLDEGSVLLHRTLGDKDSRVQNLASSLPLPEYCAHQKLNLASYLPLGLTLHPLE 978
QY 781 PQLWAAAYGVSPHRGHLGTKNLCVEADLVSVILVHADTPLPAWHRAQKDFLSGLDGGLWS 840
Db 979 PQLWAAAYGVNSHRGHLGTKNLCVEVSDLSILVHAAEQLPWPYRAQKDFLSGLDGGLWS 1038
QY 841 PGQVSTVWHVFRQAQDAQIRIRFLQMVCPAGAGALEPGAGSCYLDAGLRRRLREEWGVS 900
Db 1039 PGQVSTVWHVFRQAQDAQIRIRFLQMVCPAGAGTLEPGAGSCYLDAGLRRRLREEWGVS 1098
QY 901 CWTLLQAPGEAVLVPAGAPHVQVGLVSTVSVTQHFLSPETSALSACLCHQGPSLPPDCHL 960
Db 1099 CWTLLQAPGEAVLVPAGAPHVQVGLVSTISVTQHFLSPETSALSACLCHQGPSLPPDCHL 1158
QY 961 LYAQMDWAVFQAVKVAVGTLQEA 984
Db 1159 LYAQMDRAVFAVKAAGALQEA 1182

RESULT 5

US-09-287-354-5
; Sequence 5, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Rat
US-09-287-354-5

Query Match 75.9%; Score 4067; DB 4; Length 1207;
Best Local Similarity 77.7%; Pred. No. 0;
Matches 765; Conservative 65; Mismatches 146; Indels 8; Gaps 5;
QY 1 FYYKDPSTPRLAKEPLAAAEPLGFLNSGGHLQRAAGEAERPSLHORDGEMGAGROQNPCP 60
Db 232 FYHKDPNLRPAKEPLAAESGMLGLAPGGHLQQAQDAEGPSLHQDGETGAGROQNLCP 291
QY 61 LFLGQDPTVPWTSWPACPPGLVHTLGNVWAGPGDGNLGYQLGPPATPRCPSPPEPVTQ 120
Db 292 VFLGYDPTVPWTSWPACPPGLVHTLGNVWAGPGDGNLGYQLGPPATPRCPSPPEPVT 351
QY 121 CCSSYPPTKGGDLGPCGKQEGLEGAGSASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180
Db 352 CCSSHLPAKEDGPGCRKQDSPEGSSGPGSESSEERKA-GSRASPPSHHTKLKKTWLT 410
QY 181 RHSEQFECPCGCEVEERPVARLRALKRAGSPEVQAGMSPAPKRPDPFPFGTAEOGAGG 240
Db 411 RHSEQFECPCGCGKESPATGLRALKRAGSPEVQAGMSPAPKRPDPFPFGTAEOGAGG 469
QY 241 WQEVRTDSTGNKDVDSGQHDQKGPQDQASLQDPLQDIPCLALPAKLAQCQSCAQAG 300
Db 470 WQETPETSTGSK-AEAQQQEEQGRDGRIRLRRESRLEDTSCQHLGAVTQCPSCVQAAG 528

QY 301 EGGGHACHSQVRRSPLGGELQOEEDTATNSSSEEGPGSGPSRLSTGLAKHLLSGLGDR 360
Db 529 EVEILTSHSQSHKLPLEEKPLEEDSCAT--SEEGGSSPEASINKGLAKHLLSGLGDR 585
QY 361 LCRLLRREREAQAREGQGPVAVTGDSPGIPRCCSRCHHGLFNTHWRCPCRSRHLVCVAC 420
Db 586 LCRLLRKEREAQAREGQGPVAVTGDSPGIPHCCSRCHHGLFNTHWRCSCSHRHLVCVAC 645
QY 421 GRVACTGRAREKAGFOEQSABECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQQVVKFD 480
Db 646 GRIAGAGKNREKTGSREQRTDDCAQEAAGAACSLILTQFVSSQALAEELSTVMHQVWAKFD 705
QY 481 IRGHCPQADARVWAPGDAGQKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSETPAE 540
Db 706 IRGHCFQVQVAVWAPGDAGQKEPTKTPPAPQLSCNGSDNRTKDIKEETPDSETPAE 765
QY 541 DRAGRPPLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDRITNILDIIAQV 600
Db 766 DRAGRPPLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDRITNILDIIAQV 825
QY 601 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLLPPPGALLWLQEPQPCPRRGHFLFOEH 660
Db 826 ERKIQEKALGPGLRAGSGLRKGLSLPLSPVTRLSPPGALLWLQEP--PKHGFHLFOEH 883
QY 661 WRQGPVLVSGIQRTLQNLWGTEALGALGGVQVQALSPLGPPQSSLSGTTFWEGFSWPE 720
Db 884 WRQGPVLVSGIQKTLRLSLWGMEALGTLGGVQVQTLTALGPPQPTNLDSTAFWEGFSHPE 943
QY 721 LRPKSDGVSLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 780
Db 944 ARPKLDEGSVLLHRLPLGDKDERSRVENLASSLPLPEYCAHQKLNLASYLPLGLTLHPLE 1003
QY 781 PQLWAAAYGVSPHRGHLGTKNLCVEADLVSVILVHADTPLPAWHRAQKDFLSGLDGGLWS 840
Db 1004 PQLWAAAYGVNSHRGHLGTKNLCVEVSDLSILVHAAEQLPWPYRAQKDFLSGLDGGLWS 1063
QY 841 PGQVSTVWHVFRQAQDAQIRIRFLQMVCPAGAGALEPGAGSCYLDAGLRRRLREEWGVS 900
Db 1064 PGQVSTVWHVFRQAQDAQIRIRFLQMVCPAGAGTLEPGAGSCYLDAGLRRRLREEWGVS 1123
QY 901 CWTLLQAPGEAVLVPAGAPHVQVGLVSTVSVTQHFLSPETSALSACLCHQGPSLPPDCHL 960
Db 1124 CWTLLQAPGEAVLVPAGAPHVQVGLVSTISVTQHFLSPETSALSACLCHQGPSLPPDCHL 1183
QY 961 LYAQMDWAVFQAVKVAVGTLQEA 984
Db 1184 LYAQMDRAVFAVKAAGALQEA 1207

RESULT 6

US-09-252-991A-22176
; Sequence 22176, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22176
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22176

Query Match 4.1%; Score 219.5; DB 4; Length 705;

; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-3

Query Match		98.7%;	Score 5290;	DB 4;	Length 11189;	
Best Local Similarity		99.0%;	Pred. No. 0;			
Matches	974;	Conservative	1;	Mismatches	9;	
				Indels	0;	
				Gaps	0;	
QY	1	FYYKDPSIPRLAKEPLAAAE	PGLFGLNSGGHLQ	RAGEAERPSLHQRD	GEMGAGRQONPCP	60
Db	206	FYYKDPSIPRLAKEPLAAAE	PGLFGLNSGGHLQ	RAGEAERPSLHQRD	GEMGAGRQONPCP	265
QY	61	LFLGQPDTPWTSWPACPP	GLVHTLGNVWAGP	GDGNLGYQLGPP	ATPRCPSPPEPPVTQ	120
Db	266	LFLGQPDTPWTSWPACPP	GLVHTLGNVWAGP	GDGNLGYQLGPP	ATPRCPSPPEPPVTQ	325
QY	121	CCSSYPPTKGGDLGPCGK	CQEGLEGASGASE	PSEEVNKASGRAC	PPSHHTKLKKTWLT	180
Db	326	CCSSYPPTKGGDLGPCGK	CQEGLEGASGASE	PSEEVNKASGRAC	PPSHHTKLKKTWLT	385
QY	181	RHSEQFECPRGCPEVEER	PVARLRALKRAGS	PEVOGAMGSPAPKR	PPDPFFPGTAEQGAGG	240
Db	386	RHSEQFECPRGCPEVEER	PVARLRALKRAGS	PEVOGAMGSPAPKR	PPDPFFPGTAEQGAGG	445
QY	241	WQEVRTSIGNKDVDSG	QHDEQKGPQDGQAS	LQDPGLQDIPCLALPAKLA	QCQSCAQAAAG	300
Db	446	LQEVRTSIGNKDVDSG	QHDEQKGPQDGQAS	LQDPGLQDIPCLALPAKLA	QCQSCAQAAAG	505
QY	301	EGGHACHSQQVRRSPL	GGELQOEEDTATNSS	SEEGPGSGPDSRLSTGLAKHLLSGLGDR		360
Db	506	EGGHACHSQQVRRSPL	GGELQOEEDTATNSS	SEEGPGSGPDSRLSTGLAKHLLSGLGDR		565
QY	361	LCRLRERREALAWAQ	REGQGPVTDSPGIPRCCSRCHHGLE	NTHWRCPRCSHRLCVAC		420
Db	566	LCRLRERREALAWAQ	REGQGPVTDSPGIPRCCSRCHHGLE	NTHWRCPRCSHRLCVAC		625
QY	421	GRVAGTGRAREKAGFQ	EQSAEECTQEAGHAACSLMLTQFVSSQALAE	LSTAMHQVVKFD		480
Db	626	GRVAGTGRAREKAGFQ	EQSAEECTQEAGHAACSLMLTQFVSSQALAE	LSTAMHQVVKFD		685
QY	481	IRGHCPQADARVWAPG	DAGQOKESTQKTPPTPOPSCNGD	THRTKSIKEETPDS	SAETPAE	540
Db	686	IRGHCPQADARVWAPG	DAGQOKESTQKTPPTPOPSCNGD	THRTKSIKEETPDS	SAETPAE	745
QY	541	DRAGRGLPCPSLCELLASTAVKLC	LGHRIHMAFAPVTPALPSDDRITN	ILDSIIAQVV		600
Db	746	DRAGRGLPCPSLCELLASTAVKLC	LGHRIHMAFAPVTPALPSDDRITN	ILDSIIAQVV		805
QY	601	ERKIQEKALGPGLRAGP	GLRKGLPLSPVRPRLPPPGALLMLQEPQCP	RRRGFHLFQEH		660
Db	806	ERKIQEKALGPGLRAGP	GLRKGLPLSPVRPRLPPPGALLMLQEPQCP	RRRGFHLFQEH		865
QY	661	WRQGPVLVSGIQT	LOGNLWGTEALGALGGQVQALSPLGP	QPSSLGSTTFWEGFSWPE		720
Db	866	WRQGPVLVSGIQT	LOGNLWGTEALGALGGQVQALSPLGP	QPSSLGSTTFWEGFSWPE		925
QY	721	LRPKSDEGSVLLH	RALGDEDTSRVENLAASLPLPEY	CALHGKINLASYLPPGLALRPLE		780
Db	926	LRPKSDEGSVLLH	RALGDEDTSRVENLAASLPLPEY	CALHGKINLASYLPPGLALRPLE		985
QY	781	PQLWAAAYGVSPHRGHLG	TKNLCVEVADLVSI	LHVADTPLPAWHRAQKDFLSGLDGEGLWS		840
Db	986	PQLWAAAYGVSPHRGHLG	TKNLCVEVADLVSI	LHVADTPLPAWHRAQKDFLSGLDGEGLWS		1045
QY	841	PGSQVSTVWHVFRAQDAQ	RIRRFLQMVCPAGAGALEPGAGC	SVYLDAGLRRRLREEWGVS		900
Db	1046	PGSQVSTVWHVFRAQDAQ	RIRRFLQMVCPAGAGALEPGAGC	SVYLDAGLRRRLREEWGVS		1105

QY	901	CWTL	QAPGEAVLPAG	APHQVQGLV	STVSTQHF	LSPETSALS	AQLCHQG	PSLPPDCHL	960
Db	1106	CWTL	QAPGEAVLPAG	APHQVQGLV	STVSTQHF	LSPETSALS	AQLCHQG	ASLPPDCHL	1165
QY	961	LYAQ	MDWAVFQAV	KVAVG	TLQ	EAK	984		
Db	1166	LYAQ	MDWAVFQAV	KVAVG	TLQ	EAK	1189		

RESULT 4
US-09-287-354-6
; Sequence 6, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mouse
US-09-287-354-6

Query Match		76.6%;	Score 4102.5;	DB 4;	Length 1182;			
Best Local Similarity		78.4%;	Pred. No. 0;					
Matches	771;	Conservative	62;	Mismatches	144;			
				Indels	7;			
				Gaps	5;			
QY	1	FYYKDP	SIPRLAKEPLAAAE	PGLFGLNSGGHLQ	RAGEAERPSLHQRD	GEMGAGRQONPCP	60	
Db		:		:	:	:		
206	FYHKDP	NILRPAKEPL--	AESGMLGAPGGHLQ	QACESEGPSLHQRD	GETGAGRQONLCP	263		
QY	61	LFLGQD	TVPWTSWPACPP	GLVHTLGNVWAGP	GDGNLGYQLGPP	ATPRCPSPPEPPVTQ	120	
Db		:		:	:	:		
264	VFLGYP	DTVPRAPWP	SCPPGLVHSLGNI	WAGPGNSLGYQLG	PPATPRCPSPGPT	PPGG	323	
QY	121	CCSSYP	PTKGGDLGPCGK	CQEGLEGASGASE	PSEEVNKASGPRAC	PPSHHTKLKKTWLT	180	
Db		:		:	:	:		
324	CCSHLP	AREGDLGPCRK	CQDSPEGSSGPG	GESSEERNKADS-RAC	PPSHHTKLKKTWLT	382		
QY	181	RHSEQF	ECPRGCPEVEER	PVARLRALKRAGS	PEVQAMGSPAPKR	PPDPFFPGTAEQGAGG	240	
Db		:		:	:	:		
383	RHSEQF	ECPCGCSGKE	ESSATGLRAL	KRAGSPEVQASRG	PAPKRPSHPFPGTGR	QGARA	442	
QY	241	WQEV	RDTSIGNKDV	DSGQHQDEQKGP	DQOQASLQDPGL	QDIPCLALPAKLAQC	SCAQAAAG	300
Db		:	:	:	:	:	:	
443	WQETP	ETIIGSK-AEAE	QOEEQGRPRDGR	IRLQESRLVDTSC	QHLAGVTQCQSCVQAAG	501		
QY	301	EGGHACH	SQQVRRSPLGGEL	QOEEDTATNSS	SEEGPGSGPDSRL	STGLAKHLLSGLGDR	360	
Db		:	:	:	:	:	:	
502	EVGVLT	GHSQKSR	RSPLE-EKQLE	EEEDSSATSEEGG	GGPPEASLNKGLAKHLLSGLGDR	560		
QY	361	LCRLR	ERREALAWAQ	REGQGPVTDSPGI	PRCCSRCHHGLE	NTHWRCPRCSHRLCVAC	420	
Db		:		:	:	:		
561	LCRLR	KERREALAWAQ	REGQGPAMTED	SPGIPHCCSRCHHGLE	NTHWRCSSHRLCVAC	620		
QY	421	GRVAGT	GRAREKAGFQ	EQSAEECTQEAG	HAACSLMLTQFV	SSQALAEISTAMHQV	VVKFD	480
Db		:	:	:	:	:		
621	GRIAGAK	NREKTGSQE	QEQHTDDCAQ	EAGHAACSLILTQFV	SSQALAEISTVMHQV	WAKFD	680	
QY	481	IRGHCP	QADARVWAPG	DAGQOKESTQKT	PPTPOPSCNGD	THRTKSIKEETPD	SAETPAE	540
Db		:		:	:	:		
681	IRGHCF	CQVDARVWAPG	DGGQOKEPTEK	TPTPOPSCNGD	SNRTKDIKEETPD	STESPAE	740	
QY	541	DRAGR	GPLPCPSLCELL	ASTAVKLC	LIGHRIHMAFAP	VTPALPSDDRITN	ILDSIIAQVV	600
Db		:		:	:	:		
741	DGAGR	SPLPCPSLCELL	ASTAVKLC	LIGHRIHMAFAP	VTPALPSDDRITN	ILDSIIAQVV	800	

Db 361 LCRLLRREREALAWAQREGQGPVAVTGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
QY 421 GRVAGTGRAREKAGFQEQSABECTQEAGHAACSLMLTQFVSSQALAEALSTAMHQVWVKFD 480
Db 421 GRVAGTGRAREKAGFQEQSABECTQEAGHAACSLMLTQFVSSQALAEALSTAMHQVWVKFD 480
QY 481 IRGHCPQADARVWAPGDAGQKESTQKTPPTPPQSCNGDTHRTKSIKEETPDSAETPAE 540
Db 481 IRGHCPQADARVWAPGDAGQKESTQKTPPTPPQSCNGDTHRTKSIKEETPDSAETPAE 540
QY 541 DRAGRGPLPCPSLCELLASTAVKLCIGHERIAHMAFAPVTPALPSDDRIITNILDSIIAQVV 600
Db 541 DRAGRGPLPCPSLCELLASTAVKLCIGHERIAHMAFAPVTPALPSDDRIITNILDSIIAQVV 600
QY 601 ERKIQEKALGPGLRAGPGLRKGLGLPLSPVRPRLPPPGALLWLQEPQPCPRRGFHLFQEH 660
Db 601 ERKIQEKALGPGLRAGPGLRKGLGLPLSPVRPRLPPPGALLWLQEPQPCPRRGFHLFQEH 660
QY 661 WRQGPVLVSGIORTLQGNLWGTALGALGGQVQALSPLGPPQPSLGSSTTFWEGFSWPE 720
Db 661 WRQGPVLVSGIORTLQGNLWGTALGALGGQVQALSPLGPPQPSLGSSTTFWEGFSWPE 720
QY 721 LRPKSDSGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 780
Db 721 LRPKSDSGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 780
QY 781 PQLWAAAYGVSPHRGHLGTKNLCVEVADLVSIHVADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 781 PQLWAAAYGVSPHRGHLGTKNLCVEVADLVSIHVADTPLPAWHRAQKDFLSGLDGEGLWS 840
QY 841 PGSOVSTVWHVFRQAQAQRIIRFLQMVCPAGAGALEPGAPGSCYLDAGLRRRLREEWGVS 900
Db 841 PGSOVSTVWHVFRQAQAQRIIRFLQMVCPAGAGALEPGAPGSCYLDAGLRRRLREEWGVS 900
QY 901 CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSVTOHFLSPETSALSQAQLCHQGPSLPPDCHL 960
Db 901 CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSVTOHFLSPETSALSQAQLCHQGPSLPPDCHL 960
QY 961 LYAQMDWAVFQAVKVAVGTLQEA 984
Db 961 LYAQMDWAVFQAVKVAVGTLQEA 984

RESULT 2

US-09-287-354-4
; Sequence 4, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-4

Query Match 99.7%; Score 5343; DB 4; Length 1189;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 982; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 FYYKDPSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERP SLHQD GEMGACRQONPCP 60
Db 206 FYYKDPSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERP SLHQD GEMGACRQONPCP 265
QY 61 LFLGQPDTPVWTSWPAACPPGLVHTLGNVWAGPGDNLGYQLGPPATPRCPSPPEPVTQRG 120

Db 266 LFLGQPDTPVWTSWPAACPPGLVHTLGNVWAGPGDNLGYQLGPPATPRCPSPPEPVTQRG 325
QY 121 CCSSYPPTKGGDLGPCCKCQEGLEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180
Db 326 CCSSYPPTKGGGLGPCCKCQEGLEGGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 385
QY 181 RHSEQFECPRGCPVEERPVARLRALKRAGSPEVQAGMSPAPKRPPDPFPFGTAEQAGG 240
Db 386 RHSEQFECPRGCPVEERPVARLRALKRAGSPEVQAGMSPAPKRPPDPFPFGTAEQAGG 445
QY 241 WQEVRTDSIGNKDVDSGQDEKQPGDQASLQDPGLQDIPCLALPAKLAQCQSCAAAG 300
Db 446 WQEVRTDSIGNKDVDSGQDEKQPGDQASLQDPGLQDIPCLALPAKLAQCQSCAAAG 505
QY 301 EGGGHACHSQVRRSPLGGELQOEEDTATNSSSEEGSGSPDSRLSTGLAKHLISGLGDR 360
Db 506 EGGGHACHSQVRRSPLGGELQOEEDTATNSSSEEGSGSPDSRLSTGLAKHLISGLGDR 565
QY 361 LCRLLRREREALAWAQREGQGPVAVTGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 566 LCRLLRREREALAWAQREGQGPVAVTGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 625
QY 421 GRVAGTGRAREKAGFQEQSABECTQEAGHAACSLMLTQFVSSQALAEALSTAMHQVWVKFD 480
Db 626 GRVAGTGRAREKAGFQEQSABECTQEAGHAACSLMLTQFVSSQALAEALSTAMHQVWVKFD 685
QY 481 IRGHCPQADARVWAPGDAGQKESTQKTPPTPPQSCNGDTHRTKSIKEETPDSAETPAE 540
Db 686 IRGHCPQADARVWAPGDAGQKESTQKTPPTPPQSCNGDTHRTKSIKEETPDSAETPAE 745
QY 541 DRAGRGPLPCPSLCELLASTAVKLCIGHERIAHMAFAPVTPALPSDDRIITNILDSIIAQVV 600
Db 746 DRAGRGPLPCPSLCELLASTAVKLCIGHERIAHMAFAPVTPALPSDDRIITNILDSIIAQVV 805
QY 601 ERKIQEKALGPGLRAGPGLRKGLGLPLSPVRPRLPPPGALLWLQEPQPCPRRGFHLFQEH 660
Db 806 ERKIQEKALGPGLRAGPGLRKGLGLPLSPVRPRLPPPGALLWLQEPQPCPRRGFHLFQEH 865
QY 661 WRQGPVLVSGIORTLQGNLWGTALGALGGQVQALSPLGPPQPSLGSSTTFWEGFSWPE 720
Db 866 WRQGPVLVSGIORTLQGNLWGTALGALGGQVQALSPLGPPQPSLGSSTTFWEGFSWPE 925
QY 721 LRPKSDSGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 780
Db 926 LRPKSDSGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 985
QY 781 PQLWAAAYGVSPHRGHLGTKNLCVEVADLVSIHVADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 986 PQLWAAAYGVSPHRGHLGTKNLCVEVADLVSIHVADTPLPAWHRAQKDFLSGLDGEGLWS 1045
QY 841 PGSOVSTVWHVFRQAQAQRIIRFLQMVCPAGAGALEPGAPGSCYLDAGLRRRLREEWGVS 900
Db 1046 PGSOVSTVWHVFRQAQAQRIIRFLQMVCPAGAGALEPGAPGSCYLDAGLRRRLREEWGVS 1105
QY 901 CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSVTOHFLSPETSALSQAQLCHQGPSLPPDCHL 960
Db 1106 CWTLLQAPGEAVLVPAGAPHQVQGLVSTVSVTOHFLSPETSALSQAQLCHQGPSLPPDCHL 1165
QY 961 LYAQMDWAVFQAVKVAVGTLQEA 984
Db 1166 LYAQMDWAVFQAVKVAVGTLQEA 1189

RESULT 3

US-09-287-354-3
; Sequence 3, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 10:25:04 ; Search time 23 Seconds
(without alignments)
2208.694 Million cell updates/sec

Title: US-10-024-368-2
Perfect score: 5358
Sequence: 1 FYYKDPSIPRLAKEPLAAAE.....MDWAVEQAVKVAVGTLQEA 984

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5358	100.0	984	4	US-09-287-354-2
2	5343	99.7	1189	4	US-09-287-354-4
3	5290	98.7	1189	4	US-09-287-354-3
4	4102.5	76.6	1182	4	US-09-287-354-6
5	4067	75.9	1207	4	US-09-287-354-5
6	219.5	4.1	705	4	US-09-252-991A-22176
7	184.5	3.4	1384	3	US-08-976-255-11
8	183	3.4	2294	4	US-09-252-991A-17231
9	179	3.3	605	4	US-09-252-991A-22793
10	179	3.3	917	4	US-09-252-991A-25101
11	176.5	3.3	1043	4	US-09-252-991A-28885
12	176.5	3.3	1706	4	US-09-252-991A-31760
13	172.5	3.2	841	4	US-09-252-991A-26919
14	171.5	3.2	1149	4	US-09-252-991A-25557
15	170	3.2	1031	4	US-09-252-991A-18365
16	168	3.1	553	4	US-09-252-991A-32621
17	168	3.1	780	4	US-09-252-991A-18846
18	168	3.1	1317	3	US-09-083-521-7
19	167	3.1	822	4	US-09-252-991A-22479
20	166.5	3.1	1002	4	US-09-252-991A-27980
21	166	3.1	1366	4	US-09-585-887-10
22	166	3.1	1366	4	US-09-289-578-10
23	166	3.1	1694	1	US-08-494-168-2
24	164	3.1	633	4	US-09-252-991A-23497
25	163	3.0	955	4	US-09-252-991A-24254
26	163	3.0	1958	1	US-07-945-283-2
27	162	3.0	916	4	US-09-252-991A-31242

28	162	3.0	1088	3	US-09-130-242-2	Sequence 2, Appli
29	162	3.0	1366	3	US-08-963-825-19	Sequence 19, Appl
30	162	3.0	1366	4	US-09-500-811-19	Sequence 19, Appl
31	162	3.0	1366	4	US-09-570-573-19	Sequence 19, Appl
32	162	3.0	1366	4	US-09-548-608-19	Sequence 19, Appl
33	161.5	3.0	977	4	US-09-252-991A-16655	Sequence 16655, A
34	160.5	3.0	1057	3	US-08-931-820-4	Sequence 4, Appli
35	160	3.0	545	4	US-09-252-991A-30417	Sequence 30417, A
36	159.5	3.0	957	4	US-09-252-991A-20408	Sequence 20408, A
37	159	3.0	474	4	US-09-252-991A-28084	Sequence 28084, A
38	159	3.0	743	4	US-09-252-991A-28327	Sequence 28327, A
39	159	3.0	1180	4	US-09-252-991A-32464	Sequence 32464, A
40	159	3.0	1418	3	US-08-963-825-20	Sequence 20, Appl
41	159	3.0	1418	3	US-09-010-999-1	Sequence 1, Appli
42	159	3.0	1418	4	US-09-500-811-20	Sequence 20, Appl
43	159	3.0	1418	4	US-09-570-573-20	Sequence 20, Appl
44	159	3.0	1418	4	US-09-548-608-20	Sequence 20, Appl
45	158.5	3.0	488	4	US-09-252-991A-24759	Sequence 24759, A

ALIGNMENTS

RESULT 1
US-09-287-354-2
; Sequence 2, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-2

Query Match	100.0%;	Score	5358;	DB	4;	Length	984;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	984;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0;
Qy	1	FYYKDPSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERP	SLHQRDGEMGAGRQONPCP	60			
Db	1	FYYKDPSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERP	SLHQRDGEMGAGRQONPCP	60			
Qy	61	LFLGQDTPVPTWTSWPACPPGLVHTLGNVWAGDGNLGYQLGPP	PATPRCPSPEPPVTQRG	120			
Db	61	LFLGQDTPVPTWTSWPACPPGLVHTLGNVWAGDGNLGYQLGPP	PATPRCPSPEPPVTQRG	120			
Qy	121	CCSSYPPTKGGDLGPGCKQEGLEGASGASEPSEEVNKASGP	RACPPSHHTKLTWLT	180			
Db	121	CCSSYPPTKGGDLGPGCKQEGLEGASGASEPSEEVNKASGP	RACPPSHHTKLTWLT	180			
Qy	181	RHSEQFECPRGCPVEEERPVARLRALKRAGSPEVQAMGSP	APKRPDPFPGTAEQAGG	240			
Db	181	RHSEQFECPRGCPVEEERPVARLRALKRAGSPEVQAMGSP	APKRPDPFPGTAEQAGG	240			
Qy	241	WQEVDRDTSIGNKDVDSGQHQDEKGPQDQASLQDPGLQDIP	CLALPAKLAQCQSCAQAG	300			
Db	241	WQEVDRDTSIGNKDVDSGQHQDEKGPQDQASLQDPGLQDIP	CLALPAKLAQCQSCAQAG	300			
Qy	301	EGGHACHSQVRRSPLGGELQEEEDTATNSSSEEGSGPDS	RSLSTGLAKHLLSGLGDR	360			
Db	301	EGGHACHSQVRRSPLGGELQEEEDTATNSSSEEGSGPDS	RSLSTGLAKHLLSGLGDR	360			
Qy	361	LCRLRRERREALAWAQREGQGPVTDSPGIPRCCSRCHHGL	FNTHWRCPRCSHRLCVAC	420			

Db 324 CCSSYPPIKGGGLGPCGKCEGLDGVASGASEPSEEVNKASGRACPPGHHTKLKKTWLT 383
Qy 181 RHSEQPECPRGCPVEEERPVARLALKRAGSPEVQAMGSPAPKRPPDPFPGTAEQAGG 240
Db 384 RHSEQPECPRGCPAEAEERPVAQLRALKRAGSPEIQAGVGPAPKRPPDPFPGTAEQAGG 443
Qy 241 WQEVDTSIGNKDVDSGQHQDEQKGPQDGQASLQDPGLQDIPCLALPAKLAQCQSCAQAG 300
Db 444 WQEVDTSIGNKEADSGQHQDDQGRPRDQASLQDPGLQEIPLALPAKLAQCQSCAQAG 503
Qy 301 EGGHACHSQQVRRSPILGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360
Db 504 EGGPAGHFQQVRRSPILGGELQOEEDTAAANSSSEEGPGSGPDRSLSTGLSKHLLSSLGDR 563
Qy 361 LCRLRRERREALAWAQREGQGPVATGDSPIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 564 LCRLRRERREALAWAQREGQGPVATEDNPPIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 623
Qy 421 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 480
Db 624 GRVAGTGRAREKAGSREQSTEECTQEAGHAACSLTLTQFVSSQALAEELSTAMHQVWVKFD 683
Qy 481 IRGHCPQADARVWAPGDAGQKESTQKTPPTQPSCNGDTHRTKSIKEETPDSETPAE 540
Db 684 IRGHCPQADARVWAPGDAGQKESTQKTPPTQPSCNGDTHRTKSIKEETPDSETPAE 743
Qy 541 DRAGRPPLPCPSLCELLASTAVKLCIGHGHERIHMAFVTPALPDDRITNILDIIAQVV 600
Db 744 DRAGAPLPCLPCPSLCELLASTAVKLCIGHGHERIHMAFVTPALPDDRITNILDIIAQVV 803
Qy 601 ERKIQEKALGPLRAGPLRKLGLPLSPVRPRLPPPGALLWLQEPQPCRRRGHFLFQEH 660
Db 804 ERKIQEKALGPLRAGPLRKLGLPLSPVRPRLPPPGALLWLQEPQPCRRRGHFLFQEH 863
Qy 661 WRQGPVLVSGIQTTLQNLWGTEALGALGGVQVQALSPLGPPQPSLGSSTTFWEGFSWPE 720
Db 864 WRQGPVLVSGIQTTLQNLWGTEALGALGGVQVQALSPLGPPQPSLGSSTTFWEGFSWPE 923
Qy 721 LRPKDEGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGXINLASYLPPGLALRPLE 780
Db 924 LRPKDEGSVLLHRLALGDEDTSRVENLAASLPLPEYCAHRGXINLASYLPPGLALRPLE 983
Qy 781 POLWAAAYGVSPHRGLGTKNLCVEVADLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 984 POLWAAAYGVSPHRGLGTKNLCVEVADLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 1043
Qy 841 PGSQVSTVWHVFRQAQORIRRFQMVCPAGAGALEPGAPGCVLDAGLRRRLREEWGVS 900
Db 1044 PGSQVSTVWHVFRQAQORIRRFQMVCPAGAGALEPGAPGCVLDAGLRRRLREEWGVS 1103
Qy 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSQALCHQGPSLPPDCHL 960
Db 1104 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSQALCHQGPSLPPDCHL 1163
Qy 961 LYAQMDWAVFQAVKVAVGTLQEA 984
Db 1164 LYAQMDWAVFQAVKVAVGTLQEA 1187

RESULT 2
Q8HZ76
ID Q8HZ76 PRELIMINARY; PRT; 709 AA.
AC Q8HZ76;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hairless (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;

[1]
RN SEQUENCE FROM N.A.
RP Finocchiaro R., Portolano B., Damiani G., Caroli A., Budelli E.,
RA Bolla P., Pagnacco G.;
RT "The hairless (hr) gene is involved in the congenital hypotrichosis of
RT Valle del Belice sheep."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY130969; AAN05753.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000345; CytC heme BS.
DR PROSITE; PS00190; CYTOCHROME_C_1.
FT NON TER 709
SQ SEQUENCE 709 AA; 75435 MW; 838EC14E3A1B0071 CRC64;

Query Match 34.3%; Score 1837; DB 6; Length 709;
Best Local Similarity 68.2%; Pred. No. 3.1e-120;
Matches 348; Conservative 32; Mismatches 124; Indels 6; Gaps 3;

Qy 1 FYKDPSPRLAKEPLAAEPGLFGLNSGGHLQORAGEERPSLHQRDGEMGAGRQONPCP 60
Db 206 FYHKDPSILRLAKDPLATVEPGLLGSAPRGHLQRTGEVHPTLHQRDGETGVRHNPNC 265
Qy 61 LFLGQDTPVWTSWPACPPGLVHTLGNVWAGDGNLGYQLGPPATPRCPSPEPPVTQ 120
Db 266 LLLGHPTVARTPWPTCSPGLVHTLGNVWAVPGGSLGYQLGPSAATRCPSPGPTTQAG 325
Qy 121 CCSSYPPTKGGDLGPCGKCEGLGASGASEPSEEVNKASGRACPPSHHTKLKKTWLT 180
Db 326 YCSSHPPARD--DPCGQCQEDLEGATSEPSSEEAAYKTPSPRACLPSHHTKLKKTWLT 382
Qy 181 RHSEQPECPRGCPVEEERPVARLALKRAGSPEVQAMGSPAPKRPPDPFPGTAEQAGG 240
Db 383 RHSEQFCPCDSCPGEEESPAAQLPARKRSSPEVQGTASSPAAKRPTGPPFPGSVQGARG 442
Qy 241 WQEVDTSIGNKDVDSGQHQDEQKGPQDGQASLQDPGLQDIPCLALPAKLAQCQSCAQAG 300
Db 443 RQEVLDVFGNK-AETEQRDDHRGPRDDGASLQAPGQDTPGVPPVAGITQCQCGCAQAG 501
Qy 301 EGGHACHSQQVRRSPILGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360
Db 502 ETGSPAHYSQPLSRLLLGGEPQEEEDSA--ASSKEGGSGPEAGLSVGLAKHLLSALGDR 559
Qy 361 LCRLRRERREALAWAQREGQGPVATGDSPIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 560 LCRLRRERREALVWAHQEGAPVGTEDDPSLPRCCSCHHGLFNTHWRCPRCSHRLCVTC 619
Qy 421 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 480
Db 620 GRMAGAGSTREKAGSQGQSTKCGQEGAGHSAGSLMLTQFVSSQPLAELSAAMHQVWVKFD 679
Qy 481 IRGHCPQADARVWAPGDAGQKESTQKTP 510
Db 680 IRGHCPQADARVWTAGDGLQKEPTKTP 709

RESULT 3
Q863E6
ID Q863E6 PRELIMINARY; PRT; 342 AA.
AC Q863E6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hairless (Fragment).
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Fernandez A.I., Silio L., Rodriguez C., Noguera J.L., Sanchez A.,
RA Ovilo C.;
RT "Hairless gene porcine, mRNA (exons 11 to 19).";

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: June 29, 2004, 10:23:29 ; Search time 50 Seconds
(without alignments)
6209.398 Million cell updates/sec

Title: US-10-024-368-2
Perfect score: 5358
Sequence: 1 FYYKDPSPRLAKEPLAAAE.....MDWAVFQAVKVAVGTLQEA 984

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5119	95.5	1187	6 Q8WNL9	Q8wnl9 macaca mula
2	1837	34.3	709	6 Q8HZ76	Q8hz76 ovis aries
3	1622	30.3	342	6 Q863E6	Q863e6 sus scrofa
4	778	14.5	1265	4 Q9P2G7	Q9p2g7 homo sapien
5	777.5	14.5	2162	4 Q8N3U0	Q8n3u0 homo sapien
6	758	14.1	2210	4 Q7Z3E7	Q7z3e7 homo sapien
7	711.5	13.3	1417	4 Q9NYF4	Q9nyf4 homo sapien
8	711.5	13.3	1581	4 Q9BW93	Q9bw93 homo sapien
9	711.5	13.3	1787	4 Q9UPS0	Q9ups0 homo sapien
10	707	13.2	1338	4 Q9Y4C1	Q9y4c1 homo sapien
11	696.5	13.0	1212	4 Q9P0K0	Q9p0k0 homo sapien
12	641.5	12.0	749	11 Q8K2K4	Q8k2k4 mus musculus
13	639.5	11.9	744	11 Q8K2J6	Q8k2j6 mus musculus
14	599.5	11.2	759	4 Q9BVH6	Q9bv h6 homo sapien
15	560	10.5	1192	4 Q8IY08	Q8iy08 homo sapien
16	558	10.4	854	5 Q9VHC5	Q9vhc5 drosophila

17	482.5	9.0	592	11	Q8R350	Q8r350 mus musculu
18	478.5	8.9	492	11	Q8K2A5	Q8k2a5 mus musculu
19	475.5	8.9	444	11	Q8CCE2	Q8cce2 mus musculu
20	437.5	8.2	359	11	Q8C5Q9	Q8c5q9 mus musculu
21	335	6.3	851	10	Q65384	Q65384 arabidopsis
22	312.5	5.8	906	10	Q48794	Q48794 arabidopsis
23	306.5	5.7	1027	10	Q9SSE9	Q9sse9 arabidopsis
24	305.5	5.7	638	11	Q8C4I5	Q8c4i5 mus musculu
25	262.5	4.9	840	10	Q8VYB9	Q8vyb9 arabidopsis
26	247	4.6	1701	10	Q82022	Q82022 medicago tr
27	239.5	4.5	1629	10	Q04698	Q04698 pisum sativ
28	239	4.5	1641	10	Q41700	Q41700 vicia sativ
29	233.5	4.4	730	10	Q9SV29	Q9sv29 arabidopsis
30	218	4.1	194	11	Q9CU57	Q9cu57 mus musculu
31	213	4.0	927	10	Q8H1S7	Q8his7 arabidopsis
32	208	3.9	1208	4	Q9HCF8	Q9hcf8 homo sapien
33	204.5	3.8	1678	11	Q924C5	Q924c5 mus musculu
34	203	3.8	1285	4	Q9UPS7	Q9ups7 homo sapien
35	203	3.8	1409	4	Q8NFF9	Q8nff9 homo sapien
36	203	3.8	1419	4	Q8NFG0	Q8nfg0 homo sapien
37	199.5	3.7	1046	11	Q8K388	Q8k388 mus musculu
38	197.5	3.7	1580	11	Q8BLX7	Q8blx7 mus musculu
39	197.5	3.7	2282	11	Q61479	Q61479 mus musculu
40	195.5	3.6	5146	6	Q8SPM4	Q8spm4 bos taurus
41	195	3.6	2948	4	Q86WG6	Q86wg6 homo sapien
42	190	3.5	728	10	Q65408	Q65408 arabidopsis
43	189.5	3.5	949	4	Q7Z5T9	Q7z5t9 homo sapien
44	186.5	3.5	952	11	Q63560	Q63560 rattus norv
45	185.5	3.5	1637	6	Q9XSV8	Q9xsv8 bos taurus

ALIGNMENTS

RESULT 1

Q8WNL9 PRELIMINARY; PRT; 1187 AA.
AC Q8WNL9;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hairless.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Ahmad W., Ratterre M.S., Panteleyev A.A., Aita V.M., Sundberg J.P.,
RA Christiano A.M.;
RT "Papular Attrichia Resulting from Mutations in the Rhesus Monkey
RT (Macaca mulatta) hairless Gene."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF361864; AAL56245.1; -.
DR InterPro; IPR003347; TF_JmJC.
DR Pfam; PF02373; JmJC; 1.
DR SMART; SM00558; JmJC; 1.
SQ SEQUENCE 1187 AA; 127135 MW; D1CFED4AAA76C77E CRC64;

Query Match 95.5%; Score 5119; DB 6; Length 1187;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 938; Conservative 18; Mismatches 28; Indels 0; Gaps 0;

QY	1	FYYKDPSPRLAKEPLAAAEPLFGLNSGGHLQAGEAERPSLHQDGMGAGRQONPCP	60
Db	204	FYYKDPSPRLAKEPLAAAEPLFGLNSGGHLQAGEAERPSLHQDGMGAGRQONPCP	263
QY	61	LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPGDGNLGYQLGPPATPRCPSPPEPPVTORG	120
Db	264	LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPGSGSLGYQLGPPATPRCPSPPEPPITORG	323
QY	121	CCSSYPPTKGGDLGPCGCKQCEGLEGGASGASEPSEVNKASGRACPPSHHTKLKKTWLT	180

QY 481 IRGHCPCQADARVWAPGDAGQOQKESQKTPPTPQPSGNGDTHRTKSIKEETPDSAETPAE 540
Db 681 IRGHCFCQVDARVWAPGDGGQOQKEPTEKTPPTPQPSGNGDSNRTKDIKEETPDSTESPAE 740
QY 541 DRAGRGPLPCPSLCELLASTAVKLCIGHERIHMFAFAPVTPALPSDDRITNILDIIAQVV 600
Db 741 DGAGRSLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRITNILDIIAQVV 800
QY 601 ERKIQEKALGPGLRAGPGLRKGLGLSPVRPRLPPPGALLWLQEPQPCRRRGHFLFOEH 660
Db 801 ERKIQEKALGPGLRAGSGRLKGLSLSPVRTRLSPPGALLWLQEPQPCRRRGHFLFOEH 858
QY 661 WRQGPVLVSGIQRTLQGNLWTEALGALGGVQVQALSPLGPPQPSLSGSTTFWEGFSWPE 720
Db 859 WRQGPVLVSGIQKTLRLSLWMEALGTGGVQVQTLTALGPPQPTNLDSTAFWEGFSHPE 918
QY 721 LRPKSDGVSLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 780
Db 919 TRPKLDEGSVLLHRTLGDKDSRVQNLASSLPLPEYCAHQKLNLAASYLPLGLTLHPLE 978
QY 781 PQLWAAAGVSPHRGLGTKNLCVEVADLVSLVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 979 PQLWAAAGVNSHRGLGTKNLCVEVSDLSILVHAQAQLPPWTRAQKDFLSGLDGEGLWS 1038
QY 841 PGSQVSTVWHVFRQAQDAQRIIRFLQMVCPAGAGALEPGAPGSCYLDAGLRRRLREEWGVS 900
Db 1039 PGSQVSTVWHVFRQAQDAQRIIRFLQMVCPAGAGTLEPGAPGSCYLDAGLRRRLREEWGVS 1098
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTVQHFLSPETSALSQAQLCHQGPSLPPDCHL 960
Db 1099 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTVQHFLSPETSALSQAQLYHQGASLPPDHRM 1158
QY 961 LYAQMDWAVFQAVKAVGTLQEA 984
Db 1159 LYAQMDRAVFAVKAAGALQEA 1182

RESULT 2
S28499
probable finger protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Dec-2000
C;Accession: S28499
R;Hoog, C.; Schalling, M.; Grunder-Brundell, E.; Daneholt, B.
submitted to the EMBL Data Library, June 1991
A;Description: Analysis of a murine germ cell-specific transcript that encodes a putative
A;Reference number: S28499
A;Accession: S28499
A;Molecule type: mRNA
A;Residues: 1-1214 <HOO>
A;Cross-references: EMBL:X59993; NID:g57503; PIDN:CAA42610.1; PID:g57504
A;Experimental source: strain Sprague Dawley
C;Keywords: DNA binding; zinc finger

Query Match 13.0%; Score 694.5; DB 2; Length 1214;
Best Local Similarity 25.8%; Pred. No. 2.8e-33;
Matches 260; Conservative 159; Mismatches 397; Indels 192; Gaps 40;
QY 113 EPPVTQRCSSYPPTKG-GDL----GPCGKQEGLEGGAGGASEPSEEVNKASGPRACP 167
Db 244 KPDVCKAGLLSSKSSQVAGDLKILSEPKGSC-----IQPKTNTDQESRLESTP 292
QY 168 PSHHTKLKKTWLTTRHSEQFECPRG-CPEVEERPVARLRALKRAGS-----PEVQGA- 217
Db 293 QPVTGLTKECLVTKTSSKAELDNTAPELQKR-----LEHTASTPDGLSDKPEVEAGV 345
QY 218 --MGSPAPKRPPDPFPGTAEQAGGWQEVDRDTSIGNKDVDS---GQHDEKQGPQDQASL 272
Db 346 TRLNSCSEKK-----VGPSDLQS-QNLKETS-VKVDHDSCTRRSSNKTQTTPARKSVL 398
QY 273 QDPG-----LQDIPCLALPAKLAQCQSCA-----QAAGEGGGHACHSQVVR 313
Db 399 TDPDKLKLQSGEAFVQDDSCVNIVAQLPKCRECLDSLRLKDKQKDSPVFCRFFHFR 458

QY 314 RSPLG--GELQOQEDTATNSSSEBPG-----SGPDSRLSTGLAKHLLSGLGDLRC 362
Db 459 RLQFNKHGVLRVGEFLTTPNKYDSEALGLWLPLTKNVVGTD--LDT--AKYILANIGDHFC 514
QY 363 RLLRRREAL-----AWAQREGQGPVAVTGDSPGIPRCCSRCHHGLFNTHWRCPRS 413
Db 515 QMVISEKEAMSTIEPHRQVAVKR-----AV-----KGVREMDVCDTTFIFNLHWVCPRCG 564
QY 414 HRLCVACGRVAGTGRAREKAGFQESAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMH 473
Db 565 FGVCVDCYRLK-RKNCQQAAYKTFSWIRCVKSIHEPENLMPQTIIIPGKALYDVGDIVH 623
QY 474 QVWVKFDIRGHPCQADARVWAPGDAGQOQKESQKTPPTPQPSGNGDTHRTKSIKEETPD 533
Db 624 SVRAKWKIKANCPK-SNRQFKLFSKPALKEDLKQASLSGKPSLGMTVQSSPVLEPAV 682
QY 534 SAETPAEDRAGRGLPCPS-----LCELLASTAVK-----LCLGHERIHMAFAPVTP 580
Db 683 CGEAPSKPASNVKPI-CPANTSPLNLWADLTSGNVNKENKEKQLTMPILKNEIKCLPLP 741
QY 581 ALPSDDRITNILDIIAQVVE-----RKIQEKALGP---GLRAGPGL----- 619
Db 742 PLNKSSTVHTFNSTILTVPVSNNSGFLRNLNLSSTGKTENGLKNTPKILDDIFASLVQN 801
QY 620 -----RKGLGLPLSPVRPRLPPP-----GALLWLBPPQPCPRRGHFLFOEHWRQGP 666
Db 802 KTSSDLSKRPQGLTIKPSILGFTPHYWLCDNRLLCQDEN--NKSNNVNFRECKQGP 859
QY 667 VLVSGIQRTLQGNLWTEALGALGGVQVQALSPLGPPQPSLSGSTT--FWEGFSWPELRPK 724
Db 860 VMVSGVHHKLNTELWKPESPKBEFGEQEVDLVNCRTNEITGATVGDFWDGFDVNPRLK 919
QY 725 SD---EGSVLLHRLALGDED-----TSRVENLAASLPLPEYCALHGKLNLAASYLPPGLAL 776
Db 920 NEXEKEPMVLKLDWPPGEDFRDMMPSRFDLMANIPLEPYRRDGLKLNLAASRLPNYFVR 979
QY 777 RPLEPQLWAAAG-VSPHRGLGTKNLCVEVADLVSLVHADTPLPAWHRAQKDFLSGLDG 835
Db 980 PDLGPKMYNAYGLITPEDRKVGTNTLHLDVDSDAANVMVVGIPKQCEQEEVLRITQDG 1039
QY 836 -----EGLWSPGSQVSTVWHVFRQAQDAQRIIRFLQMV-----PAGAGALEP 877
Db 1040 DDELTIKRFIEGKEKPGA----LWHIYAAKDEKIREFLKKVSEEQQENPADH---DP 1092
QY 878 GAPGSCYLDAGLRRRLREEWGVSCTLLQAPGEAVLPAGAPHQVQGLVSTVSTVQHFLS 937
Db 1093 IHDQSWYLDRLSLRKRLYQYGVQGWAIQVFLGDDVVFIPAGAPHQVHNLVYSCIKVAEDFVS 1152
QY 938 PETSA----LSAQLCHQGPSLPPDCHLLYAQMDWAVFQAVKAVGTLQ 981
Db 1153 PEHVKHCFWLTQEFRLH--SQTHTNHEDKLQVKNVIYHAKVDAAVAMLK 1198

RESULT 3
D86254
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D86254
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86254
A;Status: preliminary

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 10:24:04 ; Search time 27 Seconds
(without alignments)
3505.647 Million cell updates/sec

Title: US-10-024-368-2
Perfect score: 5358
Sequence: 1 FYYKDPSPRLAKEPLAAAE.....MDWAVFQAVKAVGTLQEA 984

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4102.5	76.6	1182	2 I48378	hairless protein -
2	694.5	13.0	1214	2 S28499	probable finger pr
3	335	6.3	851	2 D86254	hypothetical prote
4	312.5	5.8	906	2 T01440	hypothetical prote
5	247	4.6	1701	2 T43213	ENBP1 protein - ba
6	239.5	4.5	1629	2 T06461	DNA-binding protei
7	239	4.5	1641	2 T10955	early nodulin bind
8	233.5	4.4	730	2 B85013	hypothetical prote
9	233.5	4.4	730	2 T10539	hypothetical prote
10	205	3.8	1616	2 I37183	gene APXL protein
11	199	3.7	1106	2 JQ0405	hypothetical 119.5
12	197.5	3.7	2282	2 T42717	DNA-binding protei
13	190.5	3.6	660	1 QQBE3	BHLF1 protein - hu
14	190	3.5	728	2 T05151	hypothetical prote
15	184.5	3.4	1207	2 T00378	KIAA0641 protein -
16	183	3.4	1446	1 A45344	immediate-early pr
17	182.5	3.4	1433	2 A46053	bullous pemphigoid
18	182.5	3.4	3942	2 T42730	Bassoon protein -
19	180	3.4	1733	1 B45344	probable nuclear a
20	179.5	3.4	924	2 S27923	gene LF3 protein -
21	178	3.3	677	2 S23296	collagen alpha 2(I
22	177.5	3.3	1736	2 T00391	hypothetical prote
23	177	3.3	711	2 S05381	VGF8a protein prec
24	176	3.3	2774	2 A43359	microtubule-associ
25	174.5	3.3	3938	2 T42761	Bassoon protein -
26	173.5	3.2	1146	2 A38587	collagen, cornea-s
27	172	3.2	950	2 F86222	hypothetical prote
28	171.5	3.2	2142	2 B35098	MHC class III hist
29	170.5	3.2	4957	2 T03455	ALR protein - huma

RESULT 1
I48378
hairless protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I48378
R;Cachon-Gonzalez, M.B.; Fenner, S.; Coffin, J.M.; Moran, C.; Best, S.; Stoye, J.P.
Proc. Natl. Acad. Sci. U.S.A. 91, 7717-7721, 1994
A;Title: Structure and expression of the hairless gene of mice.
A;Reference number: I48378; MUID:94329587; PMID:8052649
A;Accession: I48378
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1182 <RES>
A;Cross-references: EMBL:Z32675; NID:g531706; PIDN:CAA83587.1; PID:g531707

Query Match 76.6%; Score 4102.5; DB 2; Length 1182;
Best Local Similarity 78.4%; Pred. No. 2.2e-234;
Matches 771; Conservative 62; Mismatches 144; Indels 7; Gaps 5;

ALIGNMENTS

Qy 1 FYYKDPSPRLAKEPLAAAEPLGFLNLSGGHQLORAGEAERPSLHQRDGMGAGRQONPCP 60
Db 206 FHKDPNLRPAKEPL--AESGMLGAPGHLQQAQACESEGPSLHQRDGETGAGRQONLCP 263
Qy 61 LFLGQDTPVWTWPACPPGLVHTLGNVWAGPDGNLGYQLGPATPRCPSPPEPVTQRG 120
Db 264 VFLGYDTPVPRAPWPSPCPPGLVHSLGNWAGPSNSLGYQLGPATPRCPSPGPTPPGG 323
Qy 121 CCSSYPPTKGGDLGPCGKCGEGLEGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180
Db 324 CCSSHLPAREGDLGPCRKQDPSPEGGSSGPGESSEERNKADS-RACPPSHHTKLKKTWLT 382
Qy 181 RHSEQFECPRGCPVEEERPVARLRLALKRAGSPEVQMGSPAPKRPDPFPPTAEQAGG 240
Db 383 RHSEQFECPPGCGSGKEESSATGLRALKRAGSPEVQASRGAPAPKRPSPFPPTGRQGARA 442
Qy 241 WQEVRTSIGNKDVDSGQHDQKGPQDQASLQDPLQDIPCLALPAKLAQCQSCQAAG 300
Db 443 WQETPETIIGSK-AEAEQEEQGRGPRDGRIRLQESRLVDTSCQHLAGVTQCQSCVQAAG 501
Qy 301 EGGGHACHSQVRRSPPLGGELQEEEDTATNSSSEEGPGSGDPDSRLSTGLAKHLLSGLGDR 360
Db 502 EVGVLTHSQKRRSPL-EEKQLEEDSSATSEEGGGPGPEASLNKGLAKHLLSGLGDR 560
Qy 361 LCRLLRERREALAWAQREGQGPATVGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 561 LCRLLRERREALAWAQREGQGPAMTDSFGIPHCSCRHGLFNTHWRCSHCSHRLCVAC 620
Qy 421 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQQVVKFD 480
Db 621 GRIAGAGKNREKTGSEQHTDDCAQEAAGAACSLILTQFVSSQALAEELSTVMHQQVWAKFD 680

QY	301	EGGHACHSQVRRSP	LGELQOEEDTATNSSSEEGPGSPD	SRLSTGLAKHLLSGLGDR	360
Db	529	EVEILTSHSQSHKLPLEEKPLEEDSCAT--SEEGGSSPEASINKGLAKHLLSGLGDR	585		
QY	361	LCRLRREREAALAWAQREGQGP	AVTGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC	420	
Db	586	LCRLRREREAALAWAQREGQGP	AMTEDSPGIPHCCSRCHHGLFNTHWRCSHRLCVAC	645	
QY	421	GRVAGTGRAREKAGFQEQSAEECTQ	EAGHAACSLMLTQFVSSQALAEALSTAMHQVWVKFD	480	
Db	646	GRIAGAGKNREKTGSRQRTDDCAQ	EAGHAACSLILTQFVSSQALAEALSTVMHQVWAKFD	705	
QY	481	IRGHCPQADARVWAPGDAGQOKEST	QTKPTPTPOPSCNGDTHRTKSIKEETPD	SAETPAE	540
Db	706	IRGHCFQVDARVWAPGDGGQOKEPTEKT	PPAPQLSCNGDSNRTKDIKEETPD	STESP	765
QY	541	DRAGRGPLPCPSLCELLASTAVKLC	GLGHERIHMAFAPVTPALPSSDDRI	TNILDSIIAQVV	600
Db	766	DRAGRSPLPCPSLCELLASTAVKLC	GLGHERIHMAFAPVTPALPSSDDRI	TNILDSIIAQVV	825
QY	601	ERKIQEKALGPGLRAGPLRKGLPLSPVR	PLPPPGALLWLQEPQPCPRRGFHLFQEH	660	
Db	826	ERKIQEKALGPGLRAGSLRKGLSLPLSPVRT	QLSPPGALLWLQEP--PKHGRFLFQEH	883	
QY	661	WRQGPVLVSGIQTLOGNLWGTEAL	GALGGVQVQALSPLGPPQPSLSGSTTFWEGFSWPE	720	
Db	884	WRQGPVLVSGIQTKLRLSLWGMEAL	GTLLGGVQVQTLTALGPPQPTSLDSTAFWKGFSPHE	943	
QY	721	LRPKSDEGSVLLHRLALGDEDTSR	VENLAASLPLPEYCALHGKLNLA	SYLPPGLALRPLE	780
Db	944	ARPKLDEGSVLLHRLPLGDKDES	RVENLASSLPLPEYCAHQKLNLA	SYLPLGLTLHPLE	1003
QY	781	PQLWAAAYGVSPHRHLGTKNLCVE	ADLVSVLVHADTPLPAWHRAQKDFLSGLDGEGLWS	840	
Db	1004	PQLWAAAYGVNSHRHLGTKNLCVE	VDLISILVHAEALPPWYRAQKDFLSGLDGEGLWS	1063	
QY	841	PGSQVSTVWHVFRAQDAQIRIRFLQ	MVCPAGAGALEPGAGSCYLDAGLRRRLREEWGS	900	
Db	1064	PGSQSTVWHVFRAQDAQIRIRFLQ	MVCPAGAGTLEPGAGSCYLDAGLRRRLREEWGS	1123	
QY	901	CWTLQAPGEAVLPAGAPHQVQGLV	STVSTQHFSLPETSALSAQLCHQGPSLPPDCHL	960	
Db	1124	CWTLQAPGEAVLPAGAPHQVQGLV	STISVTQHFSLPETSALSAQLCHQGASLPPDHRM	1183	
QY	961	LYAQMDWAVFQAVKVAVGTLQ	EAK 984		
Db	1184	LYAQMDRAVVFQAVKVAVGTLQ	EAK 1207		
RESULT 15					
ABU70953					
ID	ABU70953	standard; protein; 179 AA.			
XX	AC	ABU70953;			
XX	DT	10-JUN-2003 (first entry)			
XX	DE	Human adipocyte Selected Interacting domain, SID, #584.			
XX	KW	Human; prey; adipocyte; SID; selected interacting domain; anorectic;			
KW	KW	antidiabetic; protein-protein interaction; diabetes;			
XX	XX	yeast 2-hybrid assay; metabolic disorder; obesity.			
OS	XX	Homo sapiens.			
XX	PN	WO200286122-A2.			
XX	PD	31-OCT-2002.			
XX	PF	14-MAR-2002; 2002WO-EP003768.			
XX	PR	14-MAR-2001; 2001US-0275734P.			

XX	PA	(HYBR-) HYBRIGENICS.
XX	PI	Legrain P, Daviet L;
XX	DR	WPI; 2003-103412/09.
XX	DR	N-PSDB; ACA57497.
PT	PT	New complex between two interacting proteins in adipocyte cells, useful for identifying selected interacting domains that modulate protein interactions, or for preventing or treating metabolic disorders such as obesity or diabetes.
XX	PS	Claim 6; Page 297; 382pp; English.
XX	CC	The invention relates to a complex between two interacting proteins in adipocyte cells, given in the specification. The proteins are identified by selecting a bait protein from a known adipocyte marker and then performing a yeast 2-hybrid selection to isolate prey proteins encoded by members of an adipocyte cDNA library. The proteins are designated SID (RTM) (selected interacting domains) proteins. Also included are a polynucleotide encoding a polypeptide in the adipocyte cells, a recombinant host cell expressing at least one of the interacting polypeptides of the complex, selecting a modulating compound in adipocyte cells, a SID (RTM) polypeptide comprising any of the 738 amino acid sequences given in the specification (including its fragment or variant), a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences given in the specification (including its fragment or variant), a vector comprising the SID (RTM) polynucleotide, a recombinant host cell comprising the vector, a protein chip comprising the polypeptides and a record comprising all or part of the data, listed in the specification. The complex, polypeptides, polynucleotides and compounds are useful for preventing or treating metabolic disorders such as obesity or diabetes. The polynucleotides are useful as probes or primers. The complex is particularly useful for identifying selected interacting domains (SID (RTM)) for screening drugs that modulate the protein interaction, thus exhibiting the therapeutic effect. The present sequence represents a SID (prey) protein of the invention
XX	SQ	Sequence 179 AA;
Query Match 18.0%; Score 965; DB 6; Length 179;		
Best Local Similarity 99.4%; Pred. No. 8.2e-65;		
Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	269	QASLQDPGLQDIPCLALPAKLAQCQSCAQAAAGEGGGHACHSQVRRSPLGGELQEEETA 328
Db	1	QASLQDPGLQDIPCLALPAKLAQCQSCAQAAAGEGGGHACHSQVRRSPLGGELQEEETA 60
QY	329	TNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDRLCRLRRERELAWAQREGQGPVATGDS 388
Db	61	TNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDRLCRLRRERELAWAQREGQGPVATGDS 120
QY	389	PGIPRCCSRCHHGLFNTHWRCPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEA 447
Db	121	PGIPRCCSRCHHGLFNTHWRCPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEA 179

Search completed: June 29, 2004, 10:26:19
Job time : 69 secs

CC quantitate cognate RNA and DNA within cells, which can be subsequently
CC used to correlate hair growth or loss with hairless expression or
CC hairless-regulated transcription. The amino acid sequence of hairless
CC antigen can be used for preparing specific binding molecules (e.g.
CC polyclonal or monoclonal antibodies) for monitoring protein expression,
CC for affinity purification and for functional studies. The human hairless
CC polynucleotide, polypeptide or specific binding molecule may be used to
CC identify and detect this genetic marker in family pedigrees, radiation
CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
CC identify, isolate and detect interacting proteins that bind the hairless
CC gene or protein. This sequence represents a mouse hairless polypeptide of
CC the invention
XX
SQ Sequence 1182 AA;

Query Match		76.6%;	Score 4102.5;	DB 6;	Length 1182;																																																					
Best Local Similarity		78.4%;	Pred. No. 3.5e-301;																																																							
Matches 771;		Conservative 62;	Mismatches 144;	Indels 7;	Gaps 5;																																																					
QY	1	FYKDP	SIPRLAKEPLAAAE	PGLFGLNSGGHLQ	RAGEAERPSLHQ	RDGMGAGRQ	QNP	60																																																		
Db	206	FYKDP	NILRPAKEPL--	AESGMLGAPGGH	LQQA	CESEGPSLHQ	RDGETG	AGRQNLCP 263																																																		
QY	61	LFLGQ	PDTPVPTSWP	ACPPGLVHTLGN	VWAGPGDGNL	GYQLG	PATPRCPS	PEPPVTQRG 120																																																		
Db	264	VFLGY	PDTPVRAPW	SCPPGLVHSLG	NIWAGPGS	NSLGYQLG	PATPRCPS	PGPPTPPGG 323																																																		
QY	121	CCSSYP	TKGGDLGPC	KGCKQEGLEG	GASGASEP	SEEVNKASG	PRACPPSH	TKLKKTWLT 180																																																		
Db	324	CCSSHL	PAREGDLGPC	RKCQDPS	EGGSSGPG	SESSEERN	KADS-RACPP	SHHTKLKKTWLT 382																																																		
QY	181	RHSEF	ECPRG	CEVEERPV	ARLRALKR	AGSPEVQ	GAMSPAPK	PPDPFP	PGTAEOGAGG 240																																																	
Db	383	RHSEF	ECPGCGS	GKEESSAT	GLRALKR	AGSPEVQ	ASRGAP	APKRP	SHPPFGTG	RQARA 442																																																
QY	241	WQEV	RTSIGNK	DVDSGQ	HDEQKGP	DQGOASLQ	DPGLQDIP	CLALPA	KLAQCQ	SCAQAAAG 300																																																
Db	443	WQET	PETIIGSK-AEAE	QEEQ	RGRDGR	IRLQESRLV	DTSCQHL	LAGVT	CQCSCVQAAG 501																																																	
QY	301	EGGH	ACHSQQ	VRRSPL	GGELQ	EEEDTAT	NSSEEG	PGSGPD	SRLSTGL	AKHLLSGLGDR 360																																																
Db	502	EVGL	TGHSQ	KRRSPL	-EEKQ	LEEDSSAT	SEEGG	PGPEAS	LNKGLAKHLLSGLGDR 560																																																	
QY	361	LCRL	LRER	REALAWA	QREGQ	PAVTGDS	PGIPR	CCSRCH	HGLFNTH	WRCPRCSHRLCVAC 420																																																
Db	561	LCRL	LKERE	ALAWA	QREGQ	PAMTDS	PGIPH	CCSRCH	HGLFNTH	WRCPSHCHRLCVAC 620																																																
QY	421	GRV	AGTRARE	KAGFQ	EQSAE	ECTQ	EAGHA	ACSLMLTQ	FVSSQ	ALAE	LSTAMHQ	VVKFD 480																																														
Db	621	GRI	AGAKN	REKTG	SQEQHT	DDCAQ	EAGHA	ACSLILTQ	FVSSQ	ALAE	LSTVMHQ	VWAKFD 680																																														
QY	481	IRGH	CFQADAR	VWAPG	DAGQ	QKESTQ	KTPPTQ	PPSCNG	DTHRTK	SIKEET	PD	SAETPAE 540																																														
Db	681	IRGH	CFQADAR	VWAPG	DGGQ	QKEPT	KTPTQ	PPSCNG	SDNRTK	DIKEET	PD	STESPAE 740																																														
QY	541	DRAG	GPLPCPS	ICELL	ASTAV	KLC	LCHERI	HMAFAP	VT	TPALP	SDDRIT	NILDSIIAQVV 600																																														
Db	741	DGAG	SP	LP	CP	SL	CELL	ASTAV	KLC	LGH	DR	IHMAFAP	VT	TPALP	SDDRIT	NILDSIIAQVV 800																																										
QY	601	ERKI	QEKAL	GPGL	RAGP	GLR	KGL	GL	EL	SP	VR	PR	LP	PP	PG	ALL	ML	QEP	Q	PC	RR	GF	HL	FQEH 660																																		
Db	801	ERKI	QEKAL	GPGL	RAGS	GLR	KGL	SL	PS	VR	TR	LS	PP	PG	ALL	ML	QEP	R	-PKH	GF	HL	FQEH 858																																				
QY	661	WRQ	QPV	LV	SGI	QRT	LOG	N	WG	TE	AL	G	AL	G	Q	VQ	AL	S	PL	G	P	P	S	S	L	G	T	T	F	W	E	G	F	S	W	P	E 720																					
Db	859	WRQ	QPV	LV	SGI	QKT	LR	LS	WG	ME	AL	G	T	L	G	G	V	Q	T	L	T	A	L	G	P	P	T	N	L	D	S	T	A	F	W	E	G	F	S	H	P	E 918																
QY	721	LRPK	S	D	E	G	S	V	L	L	H	R	A	L	G	D	E	D	T	S	R	V	E	N	L	A	S	L	P	L	P	E	Y	C	A	L	H	G	K	N	L	A	S	L	P	P	G	L	A	R	P	L	E 780					
Db	919	TRPK	L	D	E	G	S	V	L	L	H	R	T	L	G	D	K	A	S	R	V	Q	N	L	A	S	L	P	L	P	E	Y	C	A	H	G	K	N	L	A	S	L	P	L	L	G	L	T	L	H	L	P	L	E 978				
QY	781	PQ	L	W	A	A	Y	G	V	S	P	H	R	G	L	T	G	T	K	N	L	C	V	E	V	A	D	L	S	I	L	V	H	A	D	T	P	L	P	A	W	H	R	A	K	D	F	L	S	G	L	D	G	E	G	L	W	S 840

Db	979	PQLWAA	YGVNSH	RHGLGT	KNLCVE	VS	DLISIL	VHAEAL	PPWYRA	QKDFL	SGLD	GEGLWS	1038			
Qy	841	PGSQV	STVWHV	FRADQ	AQIRIR	FLQM	VCPAG	AGALEP	GA	PGSCY	LDAGL	RRRLRE	EWGVS 900			
Db	1039	PGSQT	STVWHV	FRADQ	AQIRIR	FLQM	VCPAG	AGATLE	PA	PGSCY	LDAGL	RRRLRE	EWGVS 1098			
Qy	901	CWTL	LQAPGE	AVLP	PAGAPH	QVGL	VSTV	TQHFL	SP	ETSA	L	SAQLY	HQ	GASL	PPD	CHL 960
Db	1099	CWTL	LQAPGE	AVLP	PAGAPH	QVGL	VSTV	TQHFL	SP	ETSA	L	SAQLY	HQ	GASL	PPD	HRM 1158
Qy	961	LYAQ	MDWAV	FQAV	KVAV	GT	LQ	EAK	984							
Db	1159	LYAQ	MDRAV	FQAV	KAAV	GT	LQ	EAK	1182							

RESULT 13
AAE19797
ID AAE19797 standard; protein; 1207 AA.
XX AAE19797;
XX AC
XX DT 18-JUN-2002 (first entry)
XX XX
DE Rat Hairless protein (Hr).
XX Rat; Hairless protein; Hr protein; HR gene; dermatological condition;
KW hair loss; gene therapy.
KW Rattus norvegicus.
OS US6348348-B1.
XX PN 19-FEB-2002.
XX PD 07-APR-1999; 99US-00287354.
XX PR 07-APR-1998; 98US-0080888P.
XX (CARN-) CARNEGIE INST WASHINGTON.
PI Thompson CC;
XX WPI; 2002-204622/26.
DR Novel expression construct, useful in the diagnosis and treatment of
XX dermatological conditions, such as hair loss, contains a Hairless gene
PT sequence.
PT Example; Fig 1; 48pp; English.
XX The invention relates to human Hairless (Hr) polypeptides and nucleic
CC acid molecules (HR) encoding such polypeptides. The invention also
CC relates to the Hairless expression constructs which may be used in
CC transcription assays. Sequences of the invention are used in diagnosis
CC and treatment of dermatological conditions such as hair loss. They are
CC also used in gene therapy. Polynucleotides of the invention can be used
CC as probes for the detection of hair loss. The present sequence is rat
XX Hairless protein (Hr)
SQ Sequence 1207 AA;

Query Match		75.9%;	Score 4067;	DB 5;	Length 1207;			
Best Local Similarity		77.7%;	Pred. No. 1.8e-298;					
Matches 765;		Conservative 65;	Mismatches 146;	Indels 8;	Gaps 5;			
QY	1	FYKDP	SIPRLAKEPLAA	BPGLFGLNSGGHLQ	RAGEAERPSLHQ	RDGMGAGRQ	QNP	60
		: : :	: : :	: : : :	: : : :	: : :	: : :	
Db	232	FYKDP	NILRPAKEPLAA	ESGMLGLAPGGHLQ	QAACDAEGPSLHQ	RDGETGAGRQ	NLCP	291
		: : :	: : :	: : : :	: : : :	: : :	: : :	
QY	61	LFLGQ	PDTPVPTSWP	ACPPGLVHTLGN	VWAGPGDGNLGY	QLGPPATPRC	PSPEPPVT	QRG 120
		: : : :	: : : :	: : :	: : :	: : : :	: : :	
Db	292	VELGY	PDTPVPTSWP	SCPPGLVHTLGN	VWAGPGS	NSFCYQLG	PPVTPRC	PSPGPPTPPG 351

QY 541 DRAGRGPLPCSLCELLASTAVKLCGLGHERIHMAFAPVTPALPSDDITNILDIIAQV 600
Db |||||||
746 DRAGRGPLPCSLCELLASTAVKLCGLGHERIHMAFAPVTPALPSDDITNILDIIAQV 805
QY 601 ERKIQEKALGPGLRAGPGLRKGLGLPLSPVRPRPLPPGALLWLQEPQPCRRGFHLFQEH 660
Db |||||||
806 ERKIQEKALGPGLRAGPGLRKGLGLPLSPVRPRPLPPGALLWLQEPQPCRRGFHLFQEH 865
QY 661 WRQGPVLVSGIORTLQGNLWGTEALGAGGVQVQALSPLGPPQPSLSGSTTFWEGFSWPE 720
Db |||||||
866 WRQGPVLVSGIORTLQGNLWGTEALGAGGVQVQALSPLGPPQPSLSGSTTFWEGFSWPE 925
QY 721 LRPKSDEGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 780
Db |||||||
926 LRPKSDEGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 985
QY 781 PQLWAAYGVSPIRHLGTLKNCVEADLVSVILVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db |||||||
986 PQLWAAYGVSPIRHLGTLKNCVEADLVSVILVHADTPLPAWHRAQKDFLSGLDGEGLWS 1045
QY 841 PGQSVTVWHVFRQAQDAQRIIRFLQVCPAGAGALEPGAPGSCVLDAGLRRRLREEWGVS 900
Db |||||||
1046 PGQSVTVWHVFRQAQDAQRIIRFLQVCPAGAGALEPGAPGSCVLDAGLRRRLREEWGVS 1105
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSQALCHQGPSLPPDCHL 960
Db |||||||
1106 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSVTQHFLSPETSALSQALCHQGPSLPPDCHL 1165
QY 961 LYAQMDWAVFQAVKVAVGTLQEA 984
Db |||||||
1166 LYAQMDWAVFQAVKVAVGTLQEA 1189

RESULT 8
AAY15218
ID AAY15218 standard; protein; 1189 AA.
XX
AC AAY15218;
XX
DT 09-NOV-1999 (first entry)
XX
DE Human Hairless mutant amino acid sequence, Thr(1022)Ala.
XX
KW alopecia; congenital alopecia; congenital atrichia;
KW androgenetic alopecia; alopecia areata; alopecia universalis; mutant;
KW hair follicle.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1022
FT /label= T1022A
FT /note= "Wild-type Thr substituted by Ala"
XX
PN WO9938965-Al.
XX
PD 05-AUG-1999.
XX
PF 29-JAN-1999; 99WO-US002128.
XX
PR 29-JAN-1998; 98US-0073043P.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Christiano AM;
XX
DR WPI; 1999-479184/40.
XX
PT Human hairless gene and protein, useful for identifying modulators of
PT hair growth.
XX
PS Claim 8; Page; 127pp; English.
XX

CC This is the amino acid sequence for the mutant human hairless protein,
CC which has Alanine substituted for Threonine at position 1022. The gene
CC was discovered by genotyping a Pakistani kindred (comprising of 4
CC affected males and 7 affected females) with an inherited form of
CC congenital alopecia universalis. The pedigree is strongly suggestive of
CC autosomal recessive inheritance. The invention provides methods and
CC sequences for the recombinant production of wild-type human hairless,
CC mutant human hairless and wild-type human whn (winged-helix-nude)
CC proteins, assays for screening for binding compounds, modulators and
CC homologues, and animal models of hairlessness. Human hairless conditions
CC such as androgenetic alopecia (male pattern baldness), alopecia areata,
CC alopecia totalis, congenital alopecia universalis, congenital alopecia
CC and severe T-cell immunodeficiency can be treated with compounds
CC identified in the assays. The methods are also useful for identifying
CC compounds that can be used to inhibit hair growth
XX

SQ Sequence 1189 AA;

Query Match 99.4%; Score 5326; DB 2; Length 1189;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 980; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 FYYKDPISIPRLAKEPLAAEPGLFGLNSGGHLQAGEAERPSLHQRDGEMGAGRQONPCP 60
Db |||||||
206 FYYKDPISIPRLAKEPLAAEPGLFGLNSGGHLQAGEAERPSLHQRDGEMGAGRQONPCP 265
QY 61 LFLGQDTPVWTWPACPPGLVHTLGNVWAGDGNLGYQLGPPATPRCPSPEPPVTQRG 120
Db |||||||
266 LFLGQDTPVWTWPACPPGLVHTLGNVWAGDGNLGYQLGPPATPRCPSPEPPVTQRG 325
QY 121 CCSSYPPTKGGDLGPCGKCGLEGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180
Db |||||||
326 CCSSYPPTKGGDLGPCGKCGLEGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 385
QY 181 RHSEQFECPRGCEVEERVARLRALKRAGSPEVQVQAMGSPAPKRPPDPFGTAEQAGG 240
Db |||||||
386 RHSEQFECPRGCEVEERVARLRALKRAGSPEVQVQAMGSPAPKRPPDPFGTAEQAGG 445
QY 241 WQEVDRDTSIGNKVDVSGQHDQKGPQDQASLQDPGLQDIPCLALPAKLAQCSCQAAG 300
Db |||||||
446 LQEVDRDTSIGNKVDVSGQHDQKGPQDQASLQDPGLQDIPCLALPAKLAQCSCQAAG 505
QY 301 EGGGHACHSQVRRSPLGGELQEEEDTATNSSSEEGSGSPDRLSTGLAKHLLSGLDR 360
Db |||||||
506 EGGGHACHSQVRRSPLGGELQEEEDTATNSSSEEGSGSPDRLSTGLAKHLLSGLDR 565
QY 361 LCRLLRREREALAWAQREGQGPVAVTGDSPGIPRCCSRCHGFLNTHWRCPRCSHRLCVAC 420
Db |||||||
566 LCRLLRREREALAWAQREGQGPVAVTGDSPGIPRCCSRCHGFLNTHWRCPRCSHRLCVAC 625
QY 421 GRVAGTGRAREKAGFQESAEECTQEAACHAACSMLTQFVSSQALAEELSTAMHQVWVKFD 480
Db |||||||
626 GRVAGTGRAREKAGFQESAEECTQEAACHAACSMLTQFVSSQALAEELSTAMHQVWVKFD 685
QY 481 IRGHCPCQADARVWAPGDAGQQKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSAEPAE 540
Db |||||||
686 IRGHCPCQADARVWAPGDAGQQKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSAEPAE 745
QY 541 DRAGRGPLPCSLCELLASTAVKLCGLGHERIHMAFAPVTPALPSDDITNILDIIAQV 600
Db |||||||
746 DRAGRGPLPCSLCELLASTAVKLCGLGHERIHMAFAPVTPALPSDDITNILDIIAQV 805
QY 601 ERKIQEKALGPGLRAGPGLRKGLGLPLSPVRPRPLPPGALLWLQEPQPCRRGFHLFQEH 660
Db |||||||
806 ERKIQEKALGPGLRAGPGLRKGLGLPLSPVRPRPLPPGALLWLQEPQPCRRGFHLFQEH 865
QY 661 WRQGPVLVSGIORTLQGNLWGTEALGAGGVQVQALSPLGPPQPSLSGSTTFWEGFSWPE 720
Db |||||||
866 WRQGPVLVSGIORTLQGNLWGTEALGAGGVQVQALSPLGPPQPSLSGSTTFWEGFSWPE 925
QY 721 LRPKSDEGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 780
Db |||||||
926 LRPKSDEGSVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 985

QY 541 DRAGRGPLPCPSLCELLASTAVKLCGHERIHMAFAPVTPALPSDDRITNILDIIAQVW 600
Db 746 DRAGRGPLPCPSLCELLASTAVKLCGHERIHMAFAPVTPALPSDDRITNILDIIAQVW 805
QY 601 ERKIQEKALGPGLRACGLRKGGLPLSPVRPRLLPPPGALLWLQEPQPCRRGFHLFQEH 660
Db 806 ERKIQEKALGPGLRACGLRKGGLPLSPVRPRLLPPPGALLWLQEPQPCRRGFHLFQEH 865
QY 661 WRQGPVLVSGIQRTLQGNLWGTEALGALGGVQVQALSPGPPQPSLSTTFWEGFSWPE 720
Db 866 WRQGPVLVSGIQRTLQGNLWGTEALGALGGVQVQALSPGPPQPSLSTTFWEGFSWPE 925
QY 721 LRPKSDGSGVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 780
Db 926 LRPKSDGSGVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLASYLPPGLALRPLE 985
QY 781 PQLWAAAYGVSPHRGHLGTKNLCVEVADLVSIIVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 986 PQLWAAAYGVSPHRGHLGTKNLCVEVADLVSIIVHADTPLPAWHRAQKDFLSGLDGEGLWS 1045
QY 841 PGSQVSTVWHVFRAQDAQRIIRFLOMVCPAGAGALEPGAGSCYLDAGLRRRLREEWGVS 900
Db 1046 PGSQVSTVWHVFRAQDAQRIIRFLOMVCPAGAGALEPGAGSCYLDAGLRRRLREEWGVS 1105
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFLSPETSALSQALCHQGPSLPDCHL 960
Db 1106 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFLSPETSALSQALCHQGPSLPDCHL 1165
QY 961 LYAQMDWAVFQAVKVAVGTLOEAK 984
Db 1166 LYAQMDWAVFQAVKVAVGTLOEAK 1189

RESULT 7
ABG72775
ID ABG72775 standard; protein; 1189 AA.
XX ABG72775;
AC
XX
DT 25-FEB-2003 (first entry)
XX Human hairless protein.
DE
XX Human; catalytic DNA; catalytic RNA; hairless protein; hair loss;
KW atrichia; hair growth; hirsutism; catalytic nucleic acid; ribozyme;
KW DNzyme; self-catalytic; hammerhead ribozyme; deoxy-ribozyme;
KW catalytic core; cleavage site; pharmaceutical; hair production;
KW hair follicle; anagen phase; catagen phase; hair removal product;
KW depilatory.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Misc-difference 446 /note= "Encoded by TGG"
FT Misc-difference 584 /note= "Encoded by GGC"
FT
XX WO200283891-A2.
PN
XX
XX 24-OCT-2002.
PD
XX 12-APR-2002; 2002WO-US011683.
PF
XX 13-APR-2001; 2001US-0283618P.
PR
XX (UYCO) UNIV COLUMBIA NEW YORK.
PA
XX Christiano AM;
XX WPI; 2003-093020/08.
PI
XX N-PSDB; ABX13987.
DR

XX New catalytic nucleic acid molecule that specifically cleaves Hairless
PT Protein mRNA, useful for inhibiting hair production by a hair-producing
PT cell, or for inhibiting transition of a hair follicle from anagen phase
PT to catagen phase.
XX Disclosure; Page 59-63; 65pp; English.
PS
XX The invention discloses a new catalytic DNA or RNA molecule that
CC specifically cleaves, or inhibits expression of, Hairless Protein mRNA
CC which comprises a catalytic domain that cleaves mRNA at a defined
CC consensus sequence and binding domains contiguous with the 5' and 3' ends
CC of the catalytic domain. Lack of expression of the hairless gene due to
CC inherited mutations leads to the complete loss of hair, known as
CC atrichia. Abundant hair growth, hirsutism, can be improved by inhibiting
CC the genes promoting hair growth, and one way to get targeted, transient
CC gene suppression is through the use of catalytic nucleic acid technology,
CC including ribozymes and DNzymes. Ribozymes are RNA structures which have
CC a self-catalytic enzymatic function and sequence specific RNA binding
CC ability. Small DNA oligonucleotides that have a similar structure to the
CC hammerhead ribozyme, called deoxy-ribozymes or DNzymes, having a
CC catalytic core and two sequence specific arms. The deoxy-ribozymes have
CC more lenient consensus cleavage site requirements and are less likely to
CC degrade, in vivo, than hammerhead ribozymes. The catalytic nucleic acids
CC are useful in pharmaceutical compositions for inhibiting hair production
CC by a hair-producing cell, for inhibiting hair growth and for inhibiting
CC the transition of a hair follicle from the anagen phase to the catagen
CC phase. A non-human transgenic mammal is useful as a model for testing
CC hair removal products which function by inhibiting hairless protein
CC expression. The sequence presented is the human hairless protein
XX
SQ Sequence 1189 AA;

Query Match 99.5%; Score 5331; DB 6; Length 1189;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 981; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 FYYKDPSSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERPSLHQDGMGAGRQONPCP 60
Db 206 FYYKDPSSIPRLAKEPLAAAEPLFGLNSGGHLQORAGEAERPSLHQDGMGAGRQONPCP 265
QY 61 LFLGQPDTPVWTWPACPPGLVHTLGNVWAGPDGDLNGLYQLGPPATPRCPSPEPPVTQRG 120
Db 266 LFLGQPDTPVWTWPACPPGLVHTLGNVWAGPDGDLNGLYQLGPPATPRCPSPEPPVTQRG 325
QY 121 CCSSYPPTKGGDLGPCGKQEGLEGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 180
Db 326 CCSSYPPTKGGDLGPCGKQEGLEGASGASEPSEEVNKASGPRACPPSHHTKLKKTWLT 385
QY 181 RHSEQFECPRGCPEVEERPVARLRALKRAGSPEVOGAMGSPAPKRPDPFPGTAEQAGG 240
Db 386 RHSEQFECPRGCPEVEERPVARLRALKRAGSPEVOGAMGSPAPKRPDPFPGTAEQAGG 445
QY 241 WQEVDRDTSIGNKDVDSGQHQDEQKGPQDQASLQDPLQDIPCLALPAKLAQCQSCAQAG 300
Db 446 LQEVDRDTSIGNKDVDSGQHQDEQKGPQDQASLQDPLQDIPCLALPAKLAQCQSCAQAG 505
QY 301 EGGGHACHSQVRRSPLGGELQOEEDTATNSSSEEGPGSGPDRLSTGLAKHLLSGLGDR 360
Db 506 EGGGHACHSQVRRSPLGGELQOEEDTATNSSSEEGPGSGPDRLSTGLAKHLLSGLGDR 565
QY 361 LCRLRRERREALAWAQREGQGPVAVTGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 566 LCRLRRERREALAWAQREGQGPVAVTGDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 625
QY 421 GRVAGTGRAREKAGFOEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQQVVKFD 480
Db 626 GRVAGTGRAREKAGFOEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQQVVKFD 685
QY 481 IRGHCPQCADARVWAPGDAGQOKESTQKTPTTPOPCSCNGDTHRTKSIKEETPDSETPAE 540
Db 686 IRGHCPQCADARVWAPGDAGQOKESTQKTPTTPOPCSCNGDTHRTKSIKEETPDSETPAE 745

QY 241 WQEVDRDTSIGNKDVDSGQHDEQKGPQDGQASLDQPLQDIPCLALPAKLAQCQSCAQAG 300
Db |||||||
241 WQEVDRDTSIGNKDVDSGQHDEQKGPQDGQASLDQPLQDIPCLALPAKLAQCQSCAQAG 300
QY 301 EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360
Db |||||||
301 EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360
QY 361 LCRLRREREALAWAQREGQGPVTDSPGIPRCCSRCHGLFNTHWRCPRCSHRLCVAC 420
Db |||||||
361 LCRLRREREALAWAQREGQGPVTDSPGIPRCCSRCHGLFNTHWRCPRCSHRLCVAC 420
QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAAGHAACSLMLTQFVSSQALAEALSTAMHQVWKFD 480
Db |||||||
421 GRVAGTGRAREKAGFQEQSAEECTQEAAGHAACSLMLTQFVSSQALAEALSTAMHQVWKFD 480
QY 481 IRGHCPQADARVWAPGDAGQKKESTQKTPPTQPPSCNGDTHRTKSIKEETPDSAETPAE 540
Db |||||||
481 IRGHCPQADARVWAPGDAGQKKESTQKTPPTQPPSCNGDTHRTKSIKEETPDSAETPAE 540
QY 541 DRAGRGPLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDRRITNILDSTIAQVV 600
Db |||||||
541 DRAGRGPLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPDDRRITNILDSTIAQVV 600
QY 601 ERKIQEKALGPGLRAGPGLRGKGLPLSPVRPRLPPPGALLWLQEPQPCPRRGHFLFQEH 660
Db |||||||
601 ERKIQEKALGPGLRAGPGLRGKGLPLSPVRPRLPPPGALLWLQEPQPCPRRGHFLFQEH 660
QY 661 WRQGPVLVSGIQTTLQGNLWTEALGALGGVQVQALSPLGPPQPSSLSGSTTFWEGFSWPE 720
Db |||||||
661 WRQGPVLVSGIQTTLQGNLWTEALGALGGVQVQALSPLGPPQPSSLSGSTTFWEGFSWPE 720
QY 721 LRPKSDGVSLLLHRLALGDEDTSRVENLAASLPLPEYCALHGKINLASYLPPGLALRPLE 780
Db |||||||
721 LRPKSDGVSLLLHRLALGDEDTSRVENLAASLPLPEYCALHGKINLASYLPPGLALRPLE 780
QY 781 PQLWAAAYGVSPHRCHLGTKNLCVEVADLVSVILVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db |||||||
781 PQLWAAAYGVSPHRCHLGTKNLCVEVADLVSVILVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
QY 841 PGSQVSTVWHVFRQAQRIIRFLQMVCPAGAGALEPGAPGSCYLDAGLRRRLREEWGVS 900
Db |||||||
841 PGSQVSTVWHVFRQAQRIIRFLQMVCPAGAGALEPGAPGSCYLDAGLRRRLREEWGVS 900
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFLSPETSALSQALCHQGPSLPPDCHL 960
Db |||||||
901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFLSPETSALSQALCHQGPSLPPDCHL 960
QY 961 LYAQMDWAVFQAVKVAVGTLQEAK 984
Db |||||||
961 LYAQMDWAVFQAVKVAVGTLQEAK 984

RESULT 6
AAY15217
ID AAY15217 standard; protein; 1189 AA.
XX
AC AAY15217;
XX
DT 09-NOV-1999 (first entry)
XX
DE Human Hairless wildtype protein amino acid sequence.
XX
KW alopecia; congenital alopecia; congenital atrichia;
KW androgenetic alopecia; alopecia areata; alopecia universalis; wildtype;
KW hair follicle.
OS Homo sapiens.
XX
PN WO9938965-Al.
XX
PD 05-AUG-1999.

XX 29-JAN-1999; 99WO-US002128.
XX
XX 29-JAN-1998; 98US-0073043P.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
XX Christiano AM;
PI
XX
DR WPI; 1999-479184/40.
DR N-PSDB; AAZ06349.
XX
PT Human hairless gene and protein, useful for identifying modulators of
hair growth.
XX
PS Claim 21; Fig 6; 127pp; English.
XX
CC This is the amino acid sequence for the wildtype human hairless protein.
CC The gene was discovered by genotyping a Pakistani kindred (comprising of
CC 4 affected males and 7 affected females) with an inherited form of
CC congenital alopecia universalis. The pedigree is strongly suggestive of
CC autosomal recessive inheritance. The invention provides methods and
CC sequences for the recombinant production of wild-type human hairless,
CC mutant human hairless and wild-type human whn (winged-helix-nude)
CC proteins, assays for screening for binding compounds, modulators and
CC homologues, and animal models of hairlessness. Human hairless conditions
CC such as androgenetic alopecia (male pattern baldness), alopecia areata,
CC alopecia totalis, congenital alopecia universalis, congenital alopecia
CC and severe T-cell immunodeficiency can be treated with compounds
CC identified in the assays. The methods are also useful for identifying
CC compounds that can be used to inhibit hair growth
XX
SQ Sequence 1189 AA;

Query Match 99.5%; Score 5331; DB 2; Length 1189;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 981; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 FYYKDPSIPRLAKEPLAAABPGLFGLNSGGHLQRAEAPRPSLHQRDGEMGAGRQNP 60
Db |||||||
206 FYYKDPSIPRLAKEPLAAABPGLFGLNSGGHLQRAEAPRPSLHQRDGEMGAGRQNP 265
QY 61 LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPDGNLQYLGPPATPRCPSPPEPVTQ 120
Db |||||||
266 LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPDGNLQYLGPPATPRCPSPPEPVTQ 325
QY 121 CCSSYPPTKGGDLGPCGKQEGLEGASGASEPSEVNVKASGPACPPSHHTKLKKTWLT 180
Db |||||||
326 CCSSYPPTKGGDLGPCGKQEGLEGASGASEPSEVNVKASGPACPPSHHTKLKKTWLT 385
QY 181 RHSEQFECPRGCPEVERPVARLRALKRAGSPEVQAMGSPAPKRPPDPFPGTAEQAGG 240
Db |||||||
386 RHSEQFECPRGCPEVERPVARLRALKRAGSPEVQAMGSPAPKRPPDPFPGTAEQAGG 445
QY 241 WQEVDRDTSIGNKDVDSGQHDEQKGPQDGQASLDQPLQDIPCLALPAKLAQCQSCAQAG 300
Db |||||||
446 LQEVDRDTSIGNKDVDSGQHDEQKGPQDGQASLDQPLQDIPCLALPAKLAQCQSCAQAG 505
QY 301 EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360
Db |||||||
506 EGGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 565
QY 361 LCRLRREREALAWAQREGQGPVTDSPGIPRCCSRCHGLFNTHWRCPRCSHRLCVAC 420
Db |||||||
566 LCRLRREREALAWAQREGQGPVTDSPGIPRCCSRCHGLFNTHWRCPRCSHRLCVAC 625
QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAAGHAACSLMLTQFVSSQALAEALSTAMHQVWKFD 480
Db |||||||
626 GRVAGTGRAREKAGFQEQSAEECTQEAAGHAACSLMLTQFVSSQALAEALSTAMHQVWKFD 685
QY 481 IRGHCPQADARVWAPGDAGQKKESTQKTPPTQPPSCNGDTHRTKSIKEETPDSAETPAE 540
Db |||||||
686 IRGHCPQADARVWAPGDAGQKKESTQKTPPTQPPSCNGDTHRTKSIKEETPDSAETPAE 745

QY 361 LCRLLRRERELAWAQREGQGPVATGDSPIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 566 LCRLLRRERELAWAQREGQGPVATEDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 625
QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 480
Db 626 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 685
QY 481 IRGHCPQADARVWAPGDAGQKKESTQKTPTTPQPSGNGDTHRTKSIKEETPDSAETPAE 540
Db 686 IRGHCPQADARVWAPGDAGQKKESTQKTPTTPQPSGNGDTHRTKSIKEETPDSAETPAE 745
QY 541 DRAGRGLPCPSLCELLASTAVKLCGHERIHMAFAPVTPALPSDDRITNILDIIIAQVV 600
Db 746 DRAGRGLPCPSLCELLASTAVKLCGHERIHMAFAPVTPALPSDDRITNILDIIIAQVV 805
QY 601 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQPCRRGFHLFQEH 660
Db 806 ERKIQEKALGPGLRAGPGLRKGLPLSPVRPRLPPPGALLWLQEPQPCRRGFHLFQEH 865
QY 661 WRQGPVLVSGIQTLOGNLWGTEALGALGGQVQALSPLGPPQPSLSGTTFWEGFSWPE 720
Db 866 WRQGPVLVSGIQTLOGNLWGTEALGALGGQVQALSPLGPPQPSLSGTTFWEGFSWPE 925
QY 721 LRPKSDGSGVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 780
Db 926 LRPKSDGSGVLLHRLALGDEDTSRVENLAASLPLPEYCALHGKLNLAASYLPPGLALRPLE 985
QY 781 PQLWAAYGVSPIRHLGTKNLCVEVADLVSIHVADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db 986 PQLWAAYGVSPIRHLGTKNLCVEVADLVSIHVADTPLPAWHRAQKDFLSGLDGEGLWS 1045
QY 841 PGSQVSTVWHVFRQAQDAQRIIRFLQMVCPAGAGALEPGAPGSCYLDAGLRRRLREEWGVS 900
Db 1046 PGSQVSTVWHVFRQAQDAQRIIRFLQMVCPAGAGALEPGAPGSCYLDAGLRRRLREEWGVS 1105
QY 901 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFSLSPETSALSQALCHQGPSLPPDCHL 960
Db 1106 CWTLLQAPGEAVLPAGAPHQVQGLVSTVSTQHFSLSPETSALSQALCHQGPSLPPDCHL 1165
QY 961 LYAQMDWAVFQAVKVAVGTILQEA 984
Db 1166 LYAQMDWAVFQAVKVAVGTILQEA 1189

RESULT 4

ABU62542
ID ABU62542 standard; protein; 1189 AA.
XX AC ABU62542;
XX 08-SEP-2003 (first entry)
DT Human Hairless (Hr) polypeptide #4.
DE Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
KW family pedigree; radiation hybrid; somatic cell hybrid.
XX Homo sapiens.
OS US2003027300-A1.
XX PN 06-FEB-2003.
XX PD 21-DEC-2001; 2001US-00024368.
XX PF 07-APR-1998; 98US-0080888P.
XX PR 07-APR-1999; 99US-00287354.
XX XX (CARN-) CARNEGIE INST WASHINGTON.
XX PA Thompson CC;
XX PI
XX XX

DR WPI; 2003-492034/58.
XX New Hairless polypeptide and polynucleotide, useful for identifying and
PT detecting this genetic marker in family pedigrees or human-rodent somatic
PT cell hybrids, or detecting interacting proteins that bind hairless gene
PT or protein.
XX Example; Fig 1; 34pp; English.
PS The invention relates to a human hairless (HR) polynucleotide encoding
XX the hairless (Hr) polypeptide. The invention also relates to an
CC expression system comprising an expression construct which produces a
CC polypeptide with hairless transcription factor activity, a reporter
CC construct comprising a transcription regulatory region responsive to
CC hairless transcription activity to regulate transcription of the reporter
CC gene which is mediated by the transcription regulatory region and methods
CC of screening for chemical agents which modulate hairless-mediated
CC transcription, binding between hairless and thyroid hormone receptor or
CC hairless activity. The polynucleotide is useful as a probe or primer to
CC quantitate cognate RNA and DNA within cells, which can be subsequently
CC used to correlate hair growth or loss with hairless expression or
CC hairless-regulated transcription. The amino acid sequence of hairless
CC antigen can be used for preparing specific binding molecules (e.g.
CC polyclonal or monoclonal antibodies) for monitoring protein expression,
CC for affinity purification and for functional studies. The human hairless
CC polynucleotide, polypeptide or specific binding molecule may be used to
CC identify and detect this genetic marker in family pedigrees, radiation
CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
CC identify, isolate and detect interacting proteins that bind the hairless
CC gene or protein. This sequence represents a human hairless polypeptide of
CC the invention
XX

SQ Sequence 1189 AA;

Query Match 99.7%; Score 5343; DB 6; Length 1189;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 982; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FYYKPSIPRLAKEPLAAEPGLFGLNSGGHLQRAEAPSLHQRDGEMGAGRQNP 60
Db 206 FYYKPSIPRLAKEPLAAEPGLFGLNSGGHLQRAEAPSLHQRDGEMGAGRQNP 265
QY 61 LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPDGDLNGLYQLGPPATPRCPSPPEPPVTQ 120
Db 266 LFLGQPDTPVWTSWPACPPGLVHTLGNVWAGPDGDLNGLYQLGPPATPRCPSPPEPPVTQ 325
QY 121 CCSSYPPTKGGDLGPCGKQEGLEGASGASEPSEEVNKASGPACPPSHHTKLKKTWLT 180
Db 326 CCSSYPPTKGGDLGPCGKQEGLEGASGASEPSEEVNKASGPACPPSHHTKLKKTWLT 385
QY 181 RHSEQFECPRGCEVEERVARLRALKRAGSPEVQAMGSPAPKRPPDPFPGTAEQAGG 240
Db 386 RHSEQFECPRGCEVEERVARLRALKRAGSPEVQAMGSPAPKRPPDPFPGTAEQAGG 445
QY 241 WQEVVDTSIGNKDVDSQHDQKGPQDGQASLDQPLQDIPCLALPAKLAQCQSCAQAG 300
Db 446 WQEVVDTSIGNKDVDSQHDQKGPQDGQASLDQPLQDIPCLALPAKLAQCQSCAQAG 505
QY 301 EGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGSGSPDSRLSTGLAKHLLSGLGDR 360
Db 506 EGGHACHSQVRRSPPLGGELQOEEDTATNSSSEEGSGSPDSRLSTGLAKHLLSGLGDR 565
QY 361 LCRLLRRERELAWAQREGQGPVATGDSPIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 420
Db 566 LCRLLRRERELAWAQREGQGPVATEDSPGIPRCCSRCHHGLFNTHWRCPRCSHRLCVAC 625
QY 421 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 480
Db 626 GRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTQFVSSQALAEELSTAMHQVWVKFD 685
QY 481 IRGHCPQADARVWAPGDAGQKKESTQKTPTTPQPSGNGDTHRTKSIKEETPDSAETPAE 540
Db 686 IRGHCPQADARVWAPGDAGQKKESTQKTPTTPQPSGNGDTHRTKSIKEETPDSAETPAE 745

Best Local Similarity 99.9%; Pred. No. 0;														
Matches 983; Conservative 0; Mismatches 1; Indels 0; Gaps 0;														
QY	1	FYYK	PSIPRLAKEPLAAAE	PGLFGLNSGGHLQ	RAGEAERPSLHQ	DGEMGAGRQ	QNP	PCP	60					
Db	1	FYYK	PSIPRLAKEPLAAAE	PGLFGLNSGGHLQ	RAGEAERPSLHQ	DGEMGAGRQ	QNP	PCP	60					
QY	61	LFLG	QPDTPVWTSWPACPP	PGLVHTLGNVWAG	PGDGNLGYQLG	PPATPRCPS	PEPPV	TQ	120					
Db	61	LFLG	QPDTPVWTSWPACPP	PGLVHTLGNVWAG	PGDGNLGYQLG	PPATPRCPS	PEPPV	TQ	120					
QY	121	CCSS	YPPTKGGDLGPCGKC	QEGLEGGASG	ASEPSEEVNKASG	PRACPPSHHT	KLKKT	WLT	180					
Db	121	CCSS	YPPTKGGDLGPCGKC	QEGLEGGASG	ASEPSEEVNKASG	PRACPPSHHT	KLKKT	WLT	180					
QY	181	RHSE	QFECPRGCPEVEER	PVARLRALKRAGS	PEVQAMGSPAP	KRPPDPFPG	TAEQ	GAGG	240					
Db	181	RHSE	QFECPRGCPEVEER	PVARLRALKRAGS	PEVQAMGSPAP	KRPPDPFPG	TAEQ	GAGG	240					
QY	241	WQEV	RDTSIGNKDVDSG	QHDQKPGQDQ	ASLQDPGLQDIP	CLALPAKLA	QCQSCA	QAAG	300					
Db	241	WQEV	RDTSIGNKDVDSG	QHDQKPGQDQ	ASLQDPGLQDIP	CLALPAKLA	QCQSCA	QAAG	300					
QY	301	EGGH	ACHSQVRRSPLG	GELQOEEDTATN	SSSEEGPGSGP	DSRLSTGL	AKHLLS	GLGDR	360					
Db	301	EGGH	ACHSQVRRSPLG	GELQOEEDTATN	SSSEEGPGSGP	DSRLSTGL	AKHLLS	GLGDR	360					
QY	361	LCRL	LRERREALAWAQ	REGQGP	AVTGDSPGIP	CRCCRHGLF	NTHWRC	PRCSHRL	CVAC	420				
Db	361	LCRL	LRERREALAWAQ	REGQGP	AVTGDSPGIP	CRCCRHGLF	NTHWRC	PRCSHRL	CVAC	420				
QY	421	GRVAG	TGRAREKAGFQ	ESAECTQEA	GHAACSLMTQ	FVSSQALAE	LSTAMH	QVWVK	FD	480				
Db	421	GRVAG	TGRAREKAGFQ	ESAECTQEA	GHAACSLMTQ	FVSSQALAE	LSTAMH	QVWVK	FD	480				
QY	481	IRGH	CPQADARVWAP	GDAQQKE	STQKTPPTQ	PSGNDTHRT	TKSIKE	ETPDS	AE	540				
Db	481	IRGH	CPQADARVWAP	GDAQQKE	STQKTPPTQ	PSGNDTHRT	TKSIKE	ETPDS	AE	540				
QY	541	DRAG	GPLPCPSLCELL	ASTAVKLC	LGHRIHMAF	AVTTPALP	SDDRIT	NILDS	IIAQ	VV	600			
Db	541	DRAG	GPLPCPSLCELL	ASTAVKLC	LGHRIHMAF	AVTTPALP	SDDRIT	NILDS	IIAQ	VV	600			
QY	601	ERKI	QEKALGPLRAG	GLKGLPLSP	VRPRLPPPG	ALLWLQEP	QPCPRR	GFHL	QEH	660				
Db	601	ERKI	QEKALGPLRAG	GLKGLPLSP	VRPRLPPPG	ALLWLQEP	QPCPRR	GFHL	QEH	660				
QY	661	WRQ	QPVLSGIQRT	LQNLWGTE	ALGALGGQ	VQALSPLG	PPQPSS	LSTTF	WEG	FSWPE	720			
Db	661	WRQ	QPVLSGIQRT	LQNLWGTE	ALGALGGQ	VQALSPLG	PPQPSS	LSTTF	WEG	FSWPE	720			
QY	721	LRPK	DEGSVLLHRL	GDDETSR	VENLAASL	PLPEYCALH	GKLNLA	SYLPP	GLALR	PLE	780			
Db	721	LRPK	DEGSVLLHRL	GDDETSR	VENLAASL	PLPEYCALH	GKLNLA	SYLPP	GLALR	PLE	780			
QY	781	PQ	LWAA	YGVSPHRG	HLGTKNLC	VEVADL	VSILVHAD	TPLPAW	HRAQK	DFLSGLD	GEG	LWS	840	
Db	781	PQ	LWAA	YGVSPHRG	HLGTKNLC	VEVADL	VSILVHAD	TPLPAW	HRAQK	DFLSGLD	GEG	LWS	840	
QY	841	PGS	QVSTVWHV	FRAQDAQ	RI	RRFLQ	MVCPAG	AGALEPG	AGSCY	LDAGL	RRRLR	EE	WGV	900
Db	841	PGS	QVSTVWHV	FRAQDAQ	RI	RRFLQ	MVCPAG	AGALEPG	AGSCY	LDAGL	RRRLR	EE	WGV	900
QY	901	CWTL	LQAPGE	AVLVPAG	APHQVQGL	VSTVS	VTQHFL	SPETS	SALSAQL	CHQGPS	LP	PD	CHL	960
Db	901	CWTL	LQAPGE	AVLVPAG	APHQVQGL	VSTVS	VTQHFL	SPETS	SALSAQL	CHQGPS	LP	PD	CHL	960
QY	961	LYAQ	MDWAVFQAV	KVAVG	T	LQ	EAK	984						
Db	961	LYAQ	MDWAVFQAV	KVAVG	T	LQ	EAK	984						

AAE19796	AAE19796 standard; protein; 1189 AA.							
AC	AAE19796;							
DT	18-JUN-2002 (first entry)							
DE	Human Hairless protein (Hr) #2.							
KW	Human; Hairless protein; Hr protein; HR gene; dermatological condition; hair loss; gene therapy.							
OS	Homo sapiens.							
PN	US6348348-B1.							
PD	19-FEB-2002.							
PF	07-APR-1999; 99US-00287354.							
PR	07-APR-1998; 98US-0080888P.							
PA	(CARN-) CARNEGIE INST WASHINGTON.							
PI	Thompson CC;							
DR	WPI; 2002-204622/26.							
XX	Novel expression construct, useful in the diagnosis and treatment of dermatological conditions, such as hair loss, contains a Hairless gene sequence.							
XX	Example; Fig 1; 48pp; English.							
CC	The invention relates to human Hairless (Hr) polypeptides and nucleic acid molecules (HR) encoding such polypeptides. The invention also relates to the Hairless expression constructs which may be used in transcription assays. Sequences of the invention are used in diagnosis and treatment of dermatological conditions such as hair loss. They are also used in gene therapy. Polynucleotides of the invention can be used as probes for the detection of hair loss. The present sequence is human Hairless protein							
SQ	Sequence 1189 AA;							
QY	Query Match 99.7%; Score 5343; DB 5; Length 1189;							
Db	Best Local Similarity 99.8%; Pred. No. 0;							
QY	Matches 982; Conservative 0; Mismatches 2; Indels 0; Gaps 0;							
1	FYYKPSIPRLAKEPLAAAE	PGLFGLNSGGHLQ	RAGEAERPSLHQ	DGEMGAGRQ	QNP	PCP	60	
206	FYYKPSIPRLAKEPLAAAE	PGLFGLNSGGHLQ	RAGEAERPSLHQ	DGEMGAGRQ	QNP	PCP	265	
61	LFLGQPDTPVWTSWPACPP	PGLVHTLGNVWAG	PGDGNLGYQLG	PPATPRCPS	PEPPV	TQ	120	
266	LFLGQPDTPVWTSWPACPP	PGLVHTLGNVWAG	PGDGNLGYQLG	PPATPRCPS	PEPPV	TQ	325	
121	CCSSYPPTKGGDLGPCGKC	QEGLEGGASG	ASEPSEEVNKASG	PRACPPSHHT	KLKKT	WLT	180	
326	CCSSYPPTKGGDLGPCGKC	QEGLEGGASG	ASEPSEEVNKASG	PRACPPSHHT	KLKKT	WLT	385	
181	RHSEQFECPRGCPEVEER	PVARLRALKRAGS	PEVQAMGSPAP	KRPPDPFPG	TAEQ	GAGG	240	
386	RHSEQFECPRGCPEVEER	PVARLRALKRAGS	PEVQAMGSPAP	KRPPDPFPG	TAEQ	GAGG	445	
241	WQEV	RDTSIGNKDVDSG	QHDQKPGQDQ	ASLQDPGLQDIP	CLALPAKLA	QCQSCA	QAAG	300
446	WQEV	RDTSIGNKDVDSG	QHDQKPGQDQ	ASLQDPGLQDIP	CLALPAKLA	QCQSCA	QAAG	505
301	EGGHACHSQVRRSPLG	GELQOEEDTATN	SSSEEGPGSGP	DSRLSTGL	AKHLLS	GLGDR	360	
506	EGGHACHSQVRRSPLG	GELQOEEDTATN	SSSEEGPGSGP	DSRLSTGL	AKHLLS	GLGDR	565	

CC	acid molecules (HR) encoding such polypeptides. The invention also
CC	relates to the Hairless expression constructs which may be used in
CC	transcription assays. Sequences of the invention are used in diagnosis
CC	and treatment of dermatological conditions such as hair loss. They are
CC	also used in gene therapy. Polynucleotides of the invention can be used
CC	as probes for the detection of hair loss. The present sequence is human
CC	Hairless protein (Hr)
XX	
SQ	Sequence 984 AA;
Query Match	
Best Local Similarity 100.0%; Score 5358; DB 5; Length 984;	
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 FYYKPSIPRLAKEPLAAAEPLFGLNSGGHLQAGEAERPSLHQRDGMGAGRQQNPCP 60
Db	
QY	1 FYYKPSIPRLAKEPLAAAEPLFGLNSGGHLQAGEAERPSLHQRDGMGAGRQQNPCP 60
Db	
QY	61 LFLGQPDTPVWTWPACPPGLVHTLGNVWAGPGDNLGYQLGPPATPRCPSPEPPVTQRG 120
Db	
QY	61 LFLGQPDTPVWTWPACPPGLVHTLGNVWAGPGDNLGYQLGPPATPRCPSPEPPVTQRG 120
Db	
QY	121 CCSSYPPTKGGDLGPCGKCGEGLGGASGASEPSEEVNKASGRACPPSHHTKLKKTWLT 180
Db	
QY	121 CCSSYPPTKGGDLGPCGKCGEGLGGASGASEPSEEVNKASGRACPPSHHTKLKKTWLT 180
Db	
QY	181 RHSEQFECPRCPEVEERPVARLRALKRAGSPEVQAMGSPAPKRPPDPFGTAEQGAGG 240
Db	
QY	181 RHSEQFECPRCPEVEERPVARLRALKRAGSPEVQAMGSPAPKRPPDPFGTAEQGAGG 240
Db	
QY	241 WQEVRTDSIGNKDVDSGGHDEQKGPQDQASLQDPLQDIPCLALPAKLAQCSCAQAG 300
Db	
QY	241 WQEVRTDSIGNKDVDSGGHDEQKGPQDQASLQDPLQDIPCLALPAKLAQCSCAQAG 300
Db	
QY	301 EGGHACHSQVRRRSPGLGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360
Db	
QY	301 EGGHACHSQVRRRSPGLGELQOEEDTATNSSSEEGPGSGPDSRLSTGLAKHLLSGLGDR 360
Db	
QY	361 LCRLRREREALAWAQREGQGPVAVTGDSPGIPRCCSRCHGLFNTHWRCPRCSHRLCVAC 420
Db	
QY	361 LCRLRREREALAWAQREGQGPVAVTGDSPGIPRCCSRCHGLFNTHWRCPRCSHRLCVAC 420
Db	
QY	421 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEELSTAMHVQVKFD 480
Db	
QY	421 GRVAGTGRAREKAGFQEQSAEECTQEAGHAACSLMLTQFVSSQALAEELSTAMHVQVKFD 480
Db	
QY	481 IRGHCPQOADARVWAPGDAGQOKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSAEPAE 540
Db	
QY	481 IRGHCPQOADARVWAPGDAGQOKESTQKTPPTPQPSGNGDTHRTKSIKEETPDSAEPAE 540
Db	
QY	541 DRAGRGLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPSDDRITNILDSTIAQVW 600
Db	
QY	541 DRAGRGLPCPSLCELLASTAVKLCGLGHERIHMAFAPVTPALPSDDRITNILDSTIAQVW 600
Db	
QY	601 ERKIQEKALGPGLRAGPGLRKGGLPLSPVRPRLPPGALLWLQEPQPCRRRGHFLFQEH 660
Db	
QY	601 ERKIQEKALGPGLRAGPGLRKGGLPLSPVRPRLPPGALLWLQEPQPCRRRGHFLFQEH 660
Db	
QY	661 WRQGPVLVSGIQTQLQNLWGTEALGALGGQVQALSPGPPQPSSLSGTTFWEGFSWPE 720
Db	
QY	661 WRQGPVLVSGIQTQLQNLWGTEALGALGGQVQALSPGPPQPSSLSGTTFWEGFSWPE 720
Db	
QY	721 LRPKSDGVSLLLHRLGDEDTSRVENLAASLPLPEYCALHGKLNLAASLPLPPGLALRPLE 780
Db	
QY	721 LRPKSDGVSLLLHRLGDEDTSRVENLAASLPLPEYCALHGKLNLAASLPLPPGLALRPLE 780
Db	
QY	781 POLWAAAYGVSPHRGLGTKNLCVEADLVSIIVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db	
QY	781 POLWAAAYGVSPHRGLGTKNLCVEADLVSIIVHADTPLPAWHRAQKDFLSGLDGEGLWS 840
Db	
QY	841 PGSQVSTVWHVFRQAQRIIRFLQMVCPAGAGALEPAGPGSCYLDAGLRRRLREEWGVS 900
Db	
QY	841 PGSQVSTVWHVFRQAQRIIRFLQMVCPAGAGALEPAGPGSCYLDAGLRRRLREEWGVS 900
Db	

QY	901	CWTLQAPGEAVLPAGAPHQVQGLVSTVSTQHFSLSPETSALSALCHQGPSLPPDCHL	960
Db	901	CWTLQAPGEAVLPAGAPHQVQGLVSTVSTQHFSLSPETSALSALCHQGPSLPPDCHL	960
QY	961	LYAQMDWAVFOAVKVAVGTLQEAK	984
Db	961	LYAQMDWAVFOAVKVAVGTLQEAK	984

RESULT 2	
ABU62545	
ID	ABU62545 standard; protein; 984 AA.
XX	
AC	ABU62545;
XX	
DT	08-SEP-2003 (first entry)
XX	
DE	Human Hairless (Hr) polypeptide #2.
XX	
KW	Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
KW	family pedigree; radiation hybrid; somatic cell hybrid.
XX	
OS	Homo sapiens.
XX	
PN	US2003027300-A1.
XX	
PD	06-FEB-2003.
XX	
PF	21-DEC-2001; 2001US-00024368.
XX	
PR	07-APR-1998; 98US-0080888P.
PR	07-APR-1999; 99US-00287354.
XX	
PA	(CARN-) CARNEGIE INST WASHINGTON.
XX	
PI	Thompson CC;
XX	
DR	WPI; 2003-492034/58.
XX	
PT	New Hairless polypeptide and polynucleotide, useful for identifying and
PT	detecting this genetic marker in family pedigrees or human-rodent somatic
PT	cell hybrids, or detecting interacting proteins that bind hairless gene
PT	or protein.
XX	

Example; Fig 1; 34pp; English.

The invention relates to a human hairless (HR) polynucleotide encoding the hairless (Hr) polypeptide. The invention also relates to an expression system comprising an expression construct which produces a polypeptide with hairless transcription factor activity, a reporter construct comprising a transcription regulatory region responsive to hairless transcription activity to regulate transcription of the reporter gene which is mediated by the transcription regulatory region and methods of screening for chemical agents which modulate hairless-mediated transcription, binding between hairless and thyroid hormone receptor or hairless activity. The polynucleotide is useful as a probe or primer to quantitate cognate RNA and DNA within cells, which can be subsequently used to correlate hair growth or loss with hairless expression or hairless-regulated transcription. The amino acid sequence of hairless antigen can be used for preparing specific binding molecules (e.g. polyclonal or monoclonal antibodies) for monitoring protein expression, for affinity purification and for functional studies. The human hairless polynucleotide, polypeptide or specific binding molecule may be used to identify and detect this genetic marker in family pedigrees, radiation hybrids or human-rodent somatic cell hybrids, and as an affinity tag to identify, isolate and detect interacting proteins that bind the hairless gene or protein. This sequence represents a human hairless polypeptide of the invention

XX SQ Sequence 984 AA;

Query Match 99.9%; Score 5353; DB 6; Length 984;

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2004, 10:19:04 ; Search time 65 Seconds
(without alignments)
4277.335 Million cell updates/sec

Title: US-10-024-368-2
Perfect score: 5358
Sequence: 1 FYYKDPISIPRLAKEPLAAAE.....MDWAVFQAVKVAVGTLQEAQ 984

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5358	100.0	984	5	AAE19794	Aae19794 Human Hai
2	5353	99.9	984	6	ABU62545	Abu62545 Human Hai
3	5343	99.7	1189	5	AAE19796	Aae19796 Human Hai
4	5343	99.7	1189	6	ABU62542	Abu62542 Human Hai
5	5334	99.6	984	6	ABU62540	Abu62540 Human Hai
6	5331	99.5	1189	2	AAI15217	Aay15217 Human Hai
7	5331	99.5	1189	6	ABG72775	Abg72775 Human hai
8	5326	99.4	1189	2	AAI15218	Aay15218 Human Hai
9	5290	98.7	1189	5	AAE19795	Aae19795 Human put
10	5290	98.7	1189	6	ABU62541	Abu62541 Human Hai
11	4102.5	76.6	1182	5	AAE19798	Aae19798 Mouse Hai
12	4102.5	76.6	1182	6	ABU62544	Abu62544 Mouse Hai
13	4067	75.9	1207	5	AAE19797	Aae19797 Rat Hairl
14	4067	75.9	1207	6	ABU62543	Abu62543 Rat Hairl
15	965	18.0	179	6	ABU70953	Abu70953 Human adi
16	777.5	14.5	1266	4	AAI40172	Aam40172 Human pol
17	777.5	14.5	2055	8	ADE40509	Ade40509 Human nuc
18	711.5	13.3	1145	3	AAI42533	Aab42533 Human ORF
19	707	13.2	1337	4	AAI78460	Aam78460 Human pro
20	707	13.2	1338	4	AAI79444	Aam79444 Human pro
21	703	13.1	1236	7	ADD18797	Add18797 Human dis
22	696.5	13.0	1212	5	ABP65070	Abp65070 Hypoxia-1
23	696.5	13.0	1213	7	ADD46413	Add46413 Human Pro
24	694.5	13.0	1214	7	ADD46411	Add46411 Rat Prote
25	574	10.7	671	5	ABP41598	Abp41598 Human ova

26	558	10.4	854	4	ABB63729	Abb63729 Drosophil
27	537	10.0	511	4	AAI41958	Aam41958 Human pol
28	278.5	5.2	837	6	ADA55130	Ada55130 Human pro
29	269.5	5.0	189	4	ABG22993	Abg22993 Novel hum
30	208	3.9	1285	5	ABG93938	Abg93938 Human pol
31	208	3.9	1400	3	AAI40638	Aab40638 Human ORF
32	208	3.9	1409	4	AAI40312	Aam40312 Human pol
33	208	3.9	1463	7	ADC39252	Adc39252 Novel hum
34	205	3.8	1646	4	AAI79494	Aam79494 Human pro
35	205	3.8	1646	7	ADE09425	Ade09425 Novel pro
36	203	3.8	1285	7	ADE73110	Ade73110 Human cel
37	203	3.8	1616	4	AAI78510	Aam78510 Human pro
38	202	3.8	900	4	ABG03533	Abg03533 Novel hum
39	200.5	3.7	19938	6	ABP76682	Abp76682 Streptomy
40	199.5	3.7	1090	7	ADC32758	Adc32758 Human nov
41	199.5	3.7	1090	7	ADC33247	Adc33247 Human nov
42	196	3.7	1564	7	ADC86801	Adc86801 Human GPC
43	194	3.6	1259	6	ABR57561	Abr57561 Human MC2
44	193.5	3.6	3640	4	ABG23029	Abg23029 Novel hum
45	193	3.6	1141	4	ABG15105	Abg15105 Novel hum

ALIGNMENTS

RESULT 1
AAE19794
ID AAE19794 standard; protein; 984 AA.
XX
AC AAE19794;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human Hairless protein (Hr) #1.
XX
KW Human; Hairless protein; Hr protein; HR gene; dermatological condition;
KW hair loss; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 284
FT /note= "Encoded by CTT of the sequence shown as SEQ ID NO:1 (AAD31072) in Column 35-38 of the specification"
FT
FT Misc-difference 978
FT /note= "Encoded by CGG of the sequence shown as SEQ ID NO:1 (AAD31077) in Column 25-32 of the specification"
FT
FT Misc-difference 982
FT /note= "Encoded by CAG of the sequence shown as SEQ ID NO:1 (AAD31077) in Column 25-32 of the specification"
FT
XX US6348348-B1.
PN
PD 19-FEB-2002.
XX
PF 07-APR-1999; 99US-00287354.
XX
PR 07-APR-1998; 98US-0080888P.
XX
PA (CARN-) CARNEGIE INST WASHINGTON.
XX
PI Thompson CC;
XX
DR WPI; 2002-204622/26.
DR N-PSDB; AAD31072, AAD31077.
XX
PT Novel expression construct, useful in the diagnosis and treatment of dermatological conditions, such as hair loss, contains a Hairless gene sequence.
XX
PS Claim 1b; Fig 1; 48pp; English.
XX
CC The invention relates to human Hairless (Hr) polypeptides and nucleic

db	5220	CCAAACACCAAGTCCACCAAGCTCACACCTGTCTTCTCAGGCTGGCATCTCCCCCAC	5279
QY	3181	CCTGTGCCCCTTTT	3193
pb	5280	CCTGTGCCCCCTTT	5292

RESULT 3

```

US-10-414-692-34
; Sequence 34, Application US/10414692
; Publication No. US20030228607A1
; GENERAL INFORMATION:
; APPLICANT: X-Ceptor Therapeutics, Inc.
; TITLE OF INVENTION: Screening method and modulators having an improved therapeutic
; TITLE OF INVENTION: profile
; FILE REFERENCE: 8012-002-US
; CURRENT APPLICATION NUMBER: US/10/414,692
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 60/372,650
; PRIOR FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 3570
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-414-692-34

```

Query Match	92.0%;	Score 2945.4;	DB 16;	Length 3570;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 2949;	Conservative	0;	Mismatches	6;
			Indels	0;
			Gaps	0;
QY	1	TTTTACTACAAGGATCCGAGCAATCCAGATTGGAAAGGAGCCCTTGGCAGCTGCGGAA	60	
Db	616	TTTTACTACAAGGATCCGAGCAATCCAGATTGGCAAGGAGCCCTTGGCAGCTGCGGAA	675	
QY	61	CCTGGTTGTTTGGCTTAAACTCTGGTGGCACTGCGAGAGCCGGGAGGCGCAACGC	120	
Db	676	CCTGGTTGTTTGGCTTAAACTCTGGTGGCACTGCGAGAGCCGGGAGGCGCAACGC	735	
QY	121	CCTTCACTGCACAGAGGGATGGAGAGATGGGAGCTGGCCGGCAGCAGAATCCTTGCCCG	180	
Db	736	CCTTCACTGCACAGAGGGATGGAGAGATGGGAGCTGGCCGGCAGCAGAATCCTTGCCCG	795	
QY	181	CTCTTCTGGGGCAGCCAGACACTGTGCCCCCTGGACCTCTGTGGCCCCGCTTGTCCCCCAGGC	240	
Db	796	CTCTTCTGGGGCAGCCAGACACTGTGCCCCCTGGACCTCTGTGGCCCCGCTTGTCCCCCAGGC	855	
QY	241	CTTGTTCAATACTCTGGCAACGCTCTGGGCTGGGCCAGGCGATGGGAACCTTGGGTACCAG	300	
Db	856	CTTGTTCAATACTCTGGCAACGCTCTGGGCTGGGCCAGGCGATGGGAACCTTGGGTACCAG	915	
QY	301	CTGGGGCCACCAAGCAACCAAGGTGCCCTCTCCTGAGCCGCTGTCAACCAGCGGGGC	360	
Db	916	CTGGGGCCACCAAGCAACCAAGGTGCCCTCTCCTGAGCCGCTGTCAACCAGCGGGGC	975	
QY	361	TGCTGTTCACTACCCACCCCACTAAAGTGGGGATCTTGGCCCTTGTGGGAAGTGCCAG	420	
Db	976	TGCTGTTCACTACCCACCCCACTAAAGTGGGGATCTTGGCCCTTGTGGGAAGTGCCAG	1035	
QY	421	GAGGGCCTGGAGGGGGTGCCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGCC	480	
Db	1036	GAGGGCCTGGAGGGGGTGCCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGCC	1095	
QY	481	TCTGGCCCCAGGSCCTGTCCCCCAGCCACACACCAAGCTGAAGAAAGACATGGCTCACA	540	
Db	1096	TCTGGCCCCAGGSCCTGTCCCCCAGCCACACACCAAGCTGAAGAAAGACATGGCTCACA	1155	
QY	541	CGGCACCTCGGAGCAATTTGAATGTCCACGCGGCTGCCCTGAGGTCGAGGAGGCGCGTT	600	
Db	1156	CGGCACCTCGGAGCAATTTGAATGTCCACGCGGCTGCCCTGAGGTCGAGGAGGCGCGTT	1215	

QY	601	GCTCGGCTCCGGGGCCCTCAAAAGCGCAGGCAGCCCGAGGTCCAGGGAGCAATGGGCAGT	660
Db	1216	GCTCGGCTCCGGGGCCCTCAAAAGCGCAGGCAGCCCGAGGTCCAGGGAGCAATGGGCAGT	1275
QY	661	CCAGCCCCAAGCGGGCCACCGGACCTTTTCCCAGGCACTGCAGAAACAGGGGGCTGGGGGT	720
Db	1276	CCAGCCCCAAGCGGGCCACCGGACCTTTTCCCAGGCACTGCAGAAACAGGGGGCTGGGGGT	1335
QY	721	TGGCAGGAGTTCGGGACACATCGATAGGGAACAAGGATGTGGACTCGGGACAGCATGAT	780
Db	1336	TGGCAGGAGTTCGGGACACATCGATAGGGAACAAGGATGTGGACTCGGGACAGCATGAT	1395
QY	781	GAGCAGAAAGAACCCCAAGATGGCCAGGCCAGTCTCCAGAACCCGGGACTTCAGGACATA	840
Db	1396	GAGCAGAAAGAACCCCAAGATGGCCAGGCCAGTCTCCAGAACCCGGGACTTCAGGACATA	1455
QY	841	CCATGCCCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCCAGGCAGCTGGA	900
Db	1456	CCATGCCCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCCAGGCAGCTGGA	1515
QY	901	GAGGAGGAGGACGCCTGCCACTCTCAGCAAGTGGGAGATCGCCTCTTGGGAGGGAG	960
Db	1516	GAGGAGGAGGACGCCTGCCACTCTCAGCAAGTGGGAGATCGCCTCTTGGGAGGGAG	1575
QY	961	CTGCAGCAGGAGGAAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCCAGGGTCCGGC	1020
Db	1576	CTGCAGCAGGAGGAAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCCAGGGTCCGGC	1635
QY	1021	CCTGACAGCCGGCTCAGCACAGGCCTCGCAAGCACCTGCTCACTGTGGTTTGGGGGACCGA	1080
Db	1636	CCTGACAGCCGGCTCAGCACAGGCCTCGCAAGCACCTGCTCACTGTGGTTTGGGGGACCGA	1695
QY	1081	CTGTGCCGCTGTCTCGGAGGGAGCGGGAGGCCCTGGCTTGGGCCAGCGGGAAGGCCAA	1140
Db	1696	CTGTGCCGCTGTCTCGGAGGGAGCGGGAGGCCCTGGCTTGGGCCAGCGGGAAGGCCAA	1755
QY	1141	GGGCCAGCCGTGACAGGGGACAGCCCAGGCATTCCACGCTGCTGCAGCCGTTGCCACCAT	1200
Db	1756	GGGCCAGCCGTGACAGGGGACAGCCCAGGCATTCCACGCTGCTGCAGCCGTTGCCACCAT	1815
QY	1201	GGACTCTTCAACACCCACTGGCGATGTCCCCTGCAGCCACCGCTGTGTGGCCCTGT	1260
Db	1816	GGACTCTTCAACACCCACTGGCGATGTCCCCTGCAGCCACCGCTGTGTGGCCCTGT	1875
QY	1261	GGTCGTGTGGCAGGCACTGGGGGGCCAGGGAAGACAGGCTTTCAGGAGCAGTCCGG	1320
Db	1876	GGTCGTGTGGCAGGCACTGGGGGGCCAGGGAAGACAGGCTTTCAGGAGCAGTCCGG	1935
QY	1321	GAGGAGTGACGCAGGAGCCGGGCACGCTGCCCTTCCCTGATGTCACCCAGTTTGT	1380
Db	1936	GAGGAGTGACGCAGGAGCCGGGCACGCTGCCCTTCCCTGATGTCACCCAGTTTGT	1995
QY	1381	TCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACCAGGTCTGGGTCAAGTTTGAT	1440
Db	1996	TCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACCAGGTCTGGGTCAAGTTTGAT	2055
QY	1441	ATCCGGGGCACTGCCCTGCCAAGCTGATGCCGGGTATGGGCCCCCGGGGATGCAGGC	1500
Db	2056	ATCCGGGGCACTGCCCTGCCAAGCTGATGCCGGGTATGGGCCCCCGGGGATGCAGGC	2115
QY	1501	CAGCAGAAGGAATCAACACAGAAAAACGCCCCCACTCCACAACCTTCTGCAATGGCGAC	1560
Db	2116	CAGCAGAAGGAATCAACACAGAAAAACGCCCCCACTCCACAACCTTCTGCAATGGCGAC	2175
QY	1561	ACCCACAGGACCAAGAGCATCAAAAGAGGAGACCCCGATTCGGCTGAGACCCAGCAGAG	1620
Db	2176	ACCCACAGGACCAAGAGCATCAAAAGAGGAGACCCCGATTCGGCTGAGACCCAGCAGAG	2235
QY	1621	GACCGTGTGGCCGAGGGCCCTTGCCTTCTCTCTGCGAACTGTGGCTTCTTACC	1680
Db	2236	GACCGTGTGGCCGAGGGCCCTTGCCTTCTCTCTGCGAACTGTGGCTTCTTACC	2295
QY	1681	GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACACATGGCCCTTCGCCCCCGTCACTCCG	1740

QY 961 CTGCAGCAGGAGGAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1020
Db |||||
QY 3060 CTGCAGCAGGAGGAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 3119
Db |||||
QY 1021 CCTGACAGCCGGCTCAGCACAGGCCTCGCCAGACCTGCTCAGTGGTTTGGGGACCGA 1080
Db |||||
QY 3120 CCTGACAGCCGGCTCAGCACAGGCCTCGCCAAAGACCTGCTCAGTGGTTTGGGGACCGA 3179
Db |||||
QY 1081 CTGTGCCGCTGTGCTCGGAGGGAGCGGGAGGCCCTGGCTTGGGCCAGCGGAAAGCCAA 1140
Db |||||
QY 3180 CTGTGCCGCTGTGCTCGGAGGGAGCGGGAGGCCCTGGCTTGGGCCAGCGGAAAGCCAA 3239
Db |||||
QY 1141 GGGCCAGCCGTGACAGGGGACAGCCAGGCATTCCACGCTGCTGACGCCGTGCCCACCAT 1200
Db |||||
QY 3240 GGGCCAGCCGTGACAGAGGACAGCCAGGCATTCCACGCTGCTGACGCCGTGCCCACCAT 3299
Db |||||
QY 1201 GGACTCTTCAACACCCACTGGCGATGTCCCGCTGCAGCCACCGGCTGTGTGGCCCTGT 1260
Db |||||
QY 3300 GGACTCTTCAACACCCACTGGCGATGTCCCGCTGCAGCCACCGGCTGTGTGGCCCTGT 3359
Db |||||
QY 1261 GGTGCTGTGCAGGCACTGGGCGGGCCAGGGAGAAAGCAGGCTTTCAGGAGCAGTCCGG 1320
Db |||||
QY 3360 GGTGCTGTGCAGGCACTGGGCGGGCCAGGGAGAAAGCAGGCTTTCAGGAGCAGTCCGG 3419
Db |||||
QY 1321 GAGGAGTGCAGCAGGAGGCGGGCAAGTGCCTGTTCCCTGATGTCACCCAGTTTGT 1380
Db |||||
QY 3420 GAGGAGTGCAGCAGGAGGCGGGCAAGTGCCTGTTCCCTGATGTCACCCAGTTTGT 3479
Db |||||
QY 1381 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACAGGCTTGGGTCAAGTTTGT 1440
Db |||||
QY 3480 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACAGGCTTGGGTCAAGTTTGT 3539
Db |||||
QY 1441 ATCCGGGGCACTGCCCTGCCAAGTGTGATGCCGGTATGGSCCCCCGGGATGCAGGC 1500
Db |||||
QY 3540 ATCCGGGGCACTGCCCTGCCAAGTGTGATGCCGGTATGGSCCCCCGGGATGCAGGC 3599
Db |||||
QY 1501 CAGCAGAAGGAATCAACACAGAAAACGCCCCAACTCCACAACTTCTGCAATGGCGAC 1560
Db |||||
QY 3600 CAGCAGAAGGAATCAACACAGAAAACGCCCCAACTCCACAACTTCTGCAATGGCGAC 3659
Db |||||
QY 1561 ACCCAGAGGACCAAGAGCATCAAGAGGAGACCCCGATTCGCTGAGACCCAGCAGAG 1620
Db |||||
QY 3660 ACCCAGAGGACCAAGAGCATCAAGAGGAGACCCCGATTCGCTGAGACCCAGCAGAG 3719
Db |||||
QY 1621 GACCGTGTGCGCGAGGGCCCTGCTGCTTCTCTCTGCGAACTGCTGGCTTCTACC 1680
Db |||||
QY 3720 GACCGTGTGCGCGAGGGCCCTGCTGCTTCTCTCTGCGAACTGCTGGCTTCTACC 3779
Db |||||
QY 1681 GCGGTCAAACTCTGCTTGGGCAATGAGCGAATACACATGGCCTTCGCCCCCGTCACTCCG 1740
Db |||||
QY 3780 GCGGTCAAACTCTGCTTGGGCAATGAGCGAATACACATGGCCTTCGCCCCCGTCACTCCG 3839
Db |||||
QY 1741 GCCCTGCCAGTGATGACCGCATCACCAACATCTTGACAGCATTATCGCACAGGTGTTG 1800
Db |||||
QY 3840 GCCCTGCCAGTGATGACCGCATCACCAACATCTTGACAGCATTATCGCACAGGTGTTG 3899
Db |||||
QY 1801 GAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGCTTCAGAGTGGCCCGGCTCTGGC 1860
Db |||||
QY 3900 GAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGCTTCAGAGTGGCCCGGCTCTGGC 3959
Db |||||
QY 1861 AAGGCCCTGGCCCTGCTCTCCAGTGGGCCCGGGCTTCAGAGTGGCCCGGCTCTGGC 1920
Db |||||
QY 3960 AAGGCCCTGGCCCTGCTCTCCAGTGGGCCCGGGCTTCAGAGTGGCCCGGCTCTGGC 4019
Db |||||
QY 1921 CTGTGGCTGCAGGAGCCCAAGCTTGCCTCGGCTGGCTTCCACCTCTTCCAGGAGCAC 1980
Db |||||
QY 4020 CTGTGGCTGCAGGAGCCCAAGCTTGCCTCGGCTGGCTTCCACCTCTTCCAGGAGCAC 4079
Db |||||
QY 1981 TGGAGGCAGGGCCAGCCTGTGTGTGAGGATCCAAAGGACATTGACAGGGCAACCTG 2040
Db |||||
QY 4080 TGGAGGCAGGGCCAGCCTGTGTGTGAGGATCCAAAGGACATTGACAGGGCAACCTG 4139
Db |||||

QY 2041 TGGGGACAGAAAGCTCTTGGGGCACTTGGAGGCCAGGTGCAGGCGTGCAGCCCCCTCGGA 2100
Db |||||
QY 4140 TGGGGACAGAAAGCTCTTGGGGCACTTGGAGGCCAGGTGCAGGCGTGCAGCCCCCTCGGA 4199
Db |||||
QY 2101 CCTCCCCAGCCAGCAGCCTGGGACGCAACAACATTTCTGGAGGGCTTCTCTGGCCTGAG 2160
Db |||||
QY 4200 CCTCCCCAGCCAGCAGCCTGGGACGCAACAACATTTCTGGAGGGCTTCTCTGGCCTGAG 4259
Db |||||
QY 2161 CTTGCCCCAAAGTCAGACGAGGGCTCTGCTCTCTCTGTCACCCAGCTTTGGGGATGAG 2220
Db |||||
QY 4260 CTTGCCCCAAAGTCAGACGAGGGCTCTGCTCTCTCTGTCACCCAGCTTTGGGGATGAG 4319
Db |||||
QY 2221 GACACCAGCAGGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCCTC 2280
Db |||||
QY 4320 GACACCAGCAGGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCCTC 4379
Db |||||
QY 2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCCTTGCCTGCGTCCACTGGAG 2340
Db |||||
QY 4380 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCCTTGCCTGCGTCCACTGGAG 4439
Db |||||
QY 2341 CCCAGCTCTGGGAGCCTATGCTGAGCCCGACCGGACCGGACACCTGGGGACCAAGAAC 2400
Db |||||
QY 4440 CCCAGCTCTGGGAGCCTATGCTGAGCCCGACCGGACCGGACACCTGGGGACCAAGAAC 4499
Db |||||
QY 2401 CTCTGTGTGGAGGTGGCCGACCTGCTCAGCATCTGCTGATGCGGACACACCACTGCCT 2460
Db |||||
QY 4500 CTCTGTGTGGAGGTGGCCGACCTGCTCAGCATCTGCTGATGCGGACACACCACTGCCT 4559
Db |||||
QY 2461 GCCTGGCACCGGGCACAGAAAGCTTCTTTCAGGCCCTGGACGGGAGGGCTCTGGTCT 2520
Db |||||
QY 4560 GCCTGGCACCGGGCACAGAAAGCTTCTTTCAGGCCCTGGACGGGAGGGCTCTGGTCT 4619
Db |||||
QY 2521 CCGGAGCAGCCAGGTGAGCACTGTGTGGACCTGTTCCGGGACAGGACGCCAGCGCATC 2580
Db |||||
QY 4620 CCGGAGCAGCCAGGTGAGCACTGTGTGGACCTGTTCCGGGACAGGACGCCAGCGCATC 4679
Db |||||
QY 2581 CGCCGCTTCTCCAGATGGTGTGCCCCGGGGCAGGCGCCTGGAGCCTGGCGCCCCA 2640
Db |||||
QY 4680 CGCCGCTTCTCCAGATGGTGTGCCCCGGGGCAGGCGCCTGGAGCCTGGCGCCCCA 4739
Db |||||
QY 2641 GGAGCTGCTACCTGGATGAGGGCTGCGGCGGCGCTTGGGGAGGAGTGGGGCGTGAGC 2700
Db |||||
QY 4740 GGAGCTGCTACCTGGATGAGGGCTGCGGCGGCGCTTGGGGAGGAGTGGGGCGTGAGC 4799
Db |||||
QY 2701 TGTGGACCTCTGCTCCAGGCCCGGAGAGGCGCTGCTGCTGCTGCGGGCTCCCCAC 2760
Db |||||
QY 4800 TGTGGACCTCTGCTCCAGGCCCGGAGAGGCGCTGCTGCTGCTGCGGGCTCCCCAC 4859
Db |||||
QY 2761 CAGGTGCAGGGCCTGCTGAGCACAGTCAGCGTCACTCAGCACTTCTCTCCCCCTGAGACC 2820
Db |||||
QY 4860 CAGGTGCAGGGCCTGCTGAGCACAGTCAGCGTCACTCAGCACTTCTCTCCCCCTGAGACC 4919
Db |||||
QY 2821 TCTGCCCTCTCTGCTCAGTCTGCCACCCAGGGACCCAGCCTTCCCCCTGACTGCCACCTG 2880
Db |||||
QY 4920 TCTGCCCTCTCTGCTCAGTCTGCCACCCAGGGACCCAGCCTTCCCCCTGACTGCCACCTG 4979
Db |||||
QY 2881 CTTTATGCCCCAGATGGACTGGGCTGTGTTCAGAGAGTGAAGGTGGCCGTGGGACATTA 2940
Db |||||
QY 4980 CTTTATGCCCCAGATGGACTGGGCTGTGTTCAGAGAGTGAAGGTGGCCGTGGGACATTA 5039
Db |||||
QY 2941 CAGGAGGCCAAATAGAGGATGCTAGGTGTCTGGATCGGGGTGGGACAGGTAGACCAG 3000
Db |||||
QY 5040 CAGGAGGCCAAATAGAGGATGCTAGGTGTCTGGATCGGGGTGGGACAGGTAGACCAG 5099
Db |||||
QY 3001 GTGCTCAGCCCCAGGACAACTTTCAGCAGGGGATGGGCTAGGGGACTTGGGATTTCTGG 3060
Db |||||
QY 5100 GTGCTCAGCCCCAGGACAACTTTCAGCAGGGGATGGGCTAGGGGACTTGGGATTTCTGG 5159
Db |||||
QY 3061 TCAACCCCCACAAGCACCCTCTGGGACAAAGCAGGGGCACTGTGTCCCCCTCCCCCTTAAG 3120
Db |||||
QY 5160 TCAACCCCCACAAGCACCCTCTGGGACAAAGCAGGGGCACTGTGTCCCCCTCCCCCTTAAG 5219
Db |||||
QY 3121 CCAACAACCACTGTCACCAAGCTCACCTGTCTTCTCAGGCTGGCATCTCCCCCAC 3180
Db |||||

QY 181 CTCTTCCTGGGAGCCAGACACACTGTGCCCCCTGGACCTCCTGGCCCCGCTTGCCCCCAGGC 240
Db |||||
181 CTCTTCCTGGGAGCCAGACACACTGTGCCCCCTGGACCTCCTGGCCCCGCTTGCCCCCAGGC 240
QY 241 CTTGTTTCATCTCTTGGCAACGTCCTGGCTGGGCGAGGCGATGGAAACCTTGGGTACCA 300
Db |||||
241 CTTGTTTCATCTCTTGGCAACGTCCTGGCTGGGCGAGGCGATGGAAACCTTGGGTACCA 300
QY 301 CTGGGCCCCACCAACAACCAAGGTGCCCCCTCTCCTGAGCCGCTGTCAACCAGCGGGC 360
Db |||||
301 CTGGGCCCCACCAACAACCAAGGTGCCCCCTCTCCTGAGCCGCTGTCAACCAGCGGGC 360
QY 361 TGCTGTTTCATCTTACCCACCCACTAAAGTGGGATCTTGGCCCTTGTGGGAAGTGCCAG 420
Db |||||
361 TGCTGTTTCATCTTACCCACCCACTAAAGTGGGATCTTGGCCCTTGTGGGAAGTGCCAG 420
QY 421 GAGGGCCTGGAGGGGGTGCAGTGGAGCCAGCAACCCAGCGAGGAAGTGAACAAGGCC 480
Db |||||
421 GAGGGCCTGGAGGGGGTGCAGTGGAGCCAGCAACCCAGCGAGGAAGTGAACAAGGCC 480
QY 481 TCTGGCCCCAGGGCCTGTCCCCCCCCAGCCACCAACCAAGCTGAAGAGACATGGCTCACA 540
Db |||||
481 TCTGGCCCCAGGGCCTGTCCCCCCCCAGCCACCAACCAAGCTGAAGAGACATGGCTCACA 540
QY 541 CGGCACTCGAGCAGTTTGAANTGCCAGCGGCTGCCCTGAGGTCGAGGAGAGGCCGGTT 600
Db |||||
541 CGGCACTCGAGCAGTTTGAANTGCCAGCGGCTGCCCTGAGGTCGAGGAGAGGCCGGTT 600
QY 601 GCTCGGTCCTGGGCCCTCAAAAGGGCAGGAGCCCGAGGTCAGGAGCAATGGGCAGT 660
Db |||||
601 GCTCGGTCCTGGGCCCTCAAAAGGGCAGGAGCCCGAGGTCAGGAGCAATGGGCAGT 660
QY 661 CCAGCCCCAAGCGGCCACCGGACCCCTTTCCAGGCACTGCAGAAACAGGGGGCTGGGGT 720
Db |||||
661 CCAGCCCCAAGCGGCCACCGGACCCCTTTCCAGGCACTGCAGAAACAGGGGGCTGGGGT 720
QY 721 TGGCAGGAGTTTCGGGACACATCGATAGGAAACAAGGATGTGACTCGGGACAGCATGAT 780
Db |||||
721 TGGCAGGAGTTTCGGGACACATCGATAGGAAACAAGGATGTGACTCGGGACAGCATGAT 780
QY 781 GAGCAGAAAGGACCCCAAGATGGCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 840
Db |||||
781 GAGCAGAAAGGACCCCAAGATGGCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 840
QY 841 CCATGCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCCAGGACGTGGA 900
Db |||||
841 CCATGCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCCAGGACGTGGA 900
QY 901 GAGGAGGAGGGCACGCTGCCACTCTCAGCAAGTGGGAGATCGCCTCTGGAGGGGAG 960
Db |||||
901 GAGGAGGAGGGCACGCTGCCACTCTCAGCAAGTGGGAGATCGCCTCTGGAGGGGAG 960
QY 961 CTGCAGCAGGAGGAAGACACAGCCCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1020
Db |||||
961 CTGCAGCAGGAGGAAGACACAGCCCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1020
QY 1021 CTTGACAGCCGGCTCAGCACAGGCTCGCCAAGCACCTGCTCAGTGGTTTGGGGACCGA 1080
Db |||||
1021 CTTGACAGCCGGCTCAGCACAGGCTCGCCAAGCACCTGCTCAGTGGTTTGGGGACCGA 1080
QY 1081 CTGTGCCGCTGTGGGAGGGAGCGGAGGCCCTGGCTTGGGCTTGGGGACCGA 1140
Db |||||
1081 CTGTGCCGCTGTGGGAGGGAGCGGAGGCCCTGGCTTGGGCTTGGGGACCGA 1140
QY 1141 GGGCAGCCGTGACAGGGGACAGCCAGGCATTCCACGCTGTGACGCCGTTGCCACAT 1200
Db |||||
1141 GGGCAGCCGTGACAGGGGACAGCCAGGCATTCCACGCTGTGACGCCGTTGCCACAT 1200
QY 1201 GGACTCTTCAACACCCACTGGGATGTCCCCGCTGCAGCCACCGGCTGTGTGGCCTGT 1260
Db |||||
1201 GGACTCTTCAACACCCACTGGGATGTCCCCGCTGCAGCCACCGGCTGTGTGGCCTGT 1260

QY 1261 GGTCTGTGGCAGGCACTGGGCGGGCCAGGAGAAAAGCAGGCTTTCAGGAGCAGTCCGCG 1320
Db |||||
1261 GGTCTGTGGCAGGCACTGGGCGGGCCAGGAGAAAAGCAGGCTTTCAGGAGCAGTCCGCG 1320
QY 1321 GAGGAGTGCACGAGGAGGCCGGCACGCTGCTGCTGCTTCCCTGATGCTGACCCAGTTTGT 1380
Db |||||
1321 GAGGAGTGCACGAGGAGGCCGGCACGCTGCTGCTGCTTCCCTGATGCTGACCCAGTTTGT 1380
QY 1381 TCCAGCCAGGCTTGGCAGAGCTGAGCACTGCAATGCAACAGGTCGAGTCAAGTTTGT 1440
Db |||||
1381 TCCAGCCAGGCTTGGCAGAGCTGAGCACTGCAATGCAACAGGTCGAGTCAAGTTTGT 1440
QY 1441 ATCCGGGGCACTGCCCTGCCAAGCTGATGCCCCGGGTATGGCCCCCGGGGATGCAGGC 1500
Db |||||
1441 ATCCGGGGCACTGCCCTGCCAAGCTGATGCCCCGGGTATGGCCCCCGGGGATGCAGGC 1500
QY 1501 CAGCAGAGGAATCAACACAGAAAACGCCCCCAACTCCACAACCTTCTGCAATGGCGAC 1560
Db |||||
1501 CAGCAGAGGAATCAACACAGAAAACGCCCCCAACTCCACAACCTTCTGCAATGGCGAC 1560
QY 1561 ACCCAGAGCAACCAAGAGCATCAAGAGAGAGACCCCCGATTCGGCTGAGACCCCGAGGAG 1620
Db |||||
1561 ACCCAGAGCAACCAAGAGCATCAAGAGAGAGACCCCCGATTCGGCTGAGACCCCGAGGAG 1620
QY 1621 GACCGTGTGGCCGAGGGCCCCCTGCTTGTCTCTCTCGAACTGCTGGCTTCTTACC 1680
Db |||||
1621 GACCGTGTGGCCGAGGGCCCCCTGCTTGTCTCTCTCGAACTGCTGGCTTCTTACC 1680
QY 1681 GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACACATGGCCCTCGCCCCCGTCACTCCG 1740
Db |||||
1681 GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACACATGGCCCTCGCCCCCGTCACTCCG 1740
QY 1741 GCCCTGCCCAGTGATGACCGCATCACCAACATCTTGACAGCATATTCGACAGGTGGTG 1800
Db |||||
1741 GCCCTGCCCAGTGATGACCGCATCACCAACATCTTGACAGCATATTCGACAGGTGGTG 1800
QY 1801 GAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGCTTCGAGCTGGCCCCGGTCTGCGC 1860
Db |||||
1801 GAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGCTTCGAGCTGGCCCCGGTCTGCGC 1860
QY 1861 AAGGCTGGGCTGCCCTCTCTCCAGTGGGCCCGGCTCGCTCCCCAGGGGCTTTG 1920
Db |||||
1861 AAGGCTGGGCTGCCCTCTCTCCAGTGGGCCCGGCTCGCTCCCCAGGGGCTTTG 1920
QY 1921 CTGTGGCTGCAGGAGCCCCAGCTTCCCTCGGCGTGGCTTCCACCTCTTCCAGGAGCAC 1980
Db |||||
1921 CTGTGGCTGCAGGAGCCCCAGCTTCCCTCGGCGTGGCTTCCACCTCTTCCAGGAGCAC 1980
QY 1981 TGGAGCAGGGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
Db |||||
1981 TGGAGCAGGGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
QY 2041 TGGGGACAGAACTCTTGGGGCACTTGGAGGCCAGGTGAGGCCCTGAGCCCCCTCGGA 2100
Db |||||
2041 TGGGGACAGAACTCTTGGGGCACTTGGAGGCCAGGTGAGGCCCTGAGCCCCCTCGGA 2100
QY 2101 CCTCCCCAGCCAGCAGCTTGGGAGCAACAATCTTGGGAGGGCTTCTCTGGGCTGAG 2160
Db |||||
2101 CCTCCCCAGCCAGCAGCTTGGGAGCAACAATCTTGGGAGGGCTTCTCTGGGCTGAG 2160
QY 2161 CTTGCCCAAAAGTCAGAGGGCTCTGTCTCTCTGCTGACCCGAGCTTGGGGATGAG 2220
Db |||||
2161 CTTGCCCAAAAGTCAGAGGGCTCTGTCTCTCTGCTGACCCGAGCTTGGGGATGAG 2220
QY 2221 GACACAGCAGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGGGCCCTC 2280
Db |||||
2221 GACACAGCAGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGGGCCCTC 2280
QY 2281 CATGAAAACTCAACTGGCTTCTTACCTCCACCGGGCTTGGCTGCTGCTGCTGCTGCTG 2340
Db |||||
2281 CATGAAAACTCAACTGGCTTCTTACCTCCACCGGGCTTGGCTGCTGCTGCTGCTGCTG 2340
QY 2341 CCCAGCTCTGGGAGCCCTATGGTGTGAGCCCGACCGGGGACACCTGGGGACCAAGAAC 2400

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 17:17:03 ; Search time 1296 Seconds
(without alignments)
11358.701 Million cell updates/sec

Title: US-10-024-368-1
Perfect score: 3202
Sequence: 1 ttttactacaaggatccgag.....tgtgcccttttatgtacagg 3202

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3031105 seqs, 2298700234 residues 6062210
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3202	100.0	3202	15	US-10-024-368-1 Sequence 1, Appli
2	3185	99.5	5709	15	US-10-122-013-1 Sequence 1, Appli
3	2945.4	92.0	3570	16	US-10-414-692-34 Sequence 34, Appli
4	1482.6	46.3	2082	17	US-10-452-858C-92 Sequence 92, Appli
5	595.8	18.6	2394	16	US-10-108-260A-2296 Sequence 2296, Ap
6	86.8	2.7	335	9	US-09-728-445-91 Sequence 91, Appli
7	66	2.1	429	9	US-09-960-352-6214 Sequence 6214, Ap
8	55.4	1.7	65	10	US-09-908-975-25105 Sequence 25105, A
9	49	1.5	125401	17	US-10-203-295-35 Sequence 35, Appli
10	48.8	1.5	9025608	15	US-10-156-761-1 Sequence 1, Appli
11	48.2	1.5	479	10	US-09-918-995-21655 Sequence 21655, A
12	48.2	1.5	484	13	US-10-085-783A-26116 Sequence 26116, A
13	48.2	1.5	484	16	US-10-242-535A-26116 Sequence 26116, A
14	48.2	1.5	3044	16	US-10-264-049-555 Sequence 555, App

15	48.2	1.5	6995	13	US-10-193-874-19	Sequence 19, Appl
16	48.2	1.5	8128	13	US-10-193-874-12	Sequence 12, Appl
17	48.2	1.5	8234	13	US-10-193-874-11	Sequence 11, Appl
18	48.2	1.5	8275	13	US-10-193-874-10	Sequence 10, Appl
19	48.2	1.5	8479	13	US-10-193-874-9	Sequence 9, Appli
20	48.2	1.5	8515	13	US-10-193-874-7	Sequence 7, Appli
21	48.2	1.5	8516	13	US-10-193-874-8	Sequence 8, Appli
22	48.2	1.5	8552	13	US-10-193-874-6	Sequence 6, Appli
23	48.2	1.5	8585	13	US-10-193-874-5	Sequence 5, Appli
24	48.2	1.5	8622	13	US-10-193-874-4	Sequence 4, Appli
25	48.2	1.5	8656	13	US-10-193-874-3	Sequence 3, Appli
26	48.2	1.5	8693	13	US-10-193-874-2	Sequence 2, Appli
27	48.2	1.5	8734	13	US-10-193-874-1	Sequence 1, Appli
28	48.2	1.5	82121	16	US-10-085-117-136	Sequence 136, App
29	47.8	1.5	1089	10	US-09-373-658-75	Sequence 75, Appl
30	47.8	1.5	1089	11	US-09-989-687-75	Sequence 75, Appl
31	47.4	1.5	651	15	US-10-267-430-11	Sequence 11, Appl
32	47.4	1.5	653	15	US-10-184-644-402	Sequence 402, App
33	47.4	1.5	653	15	US-10-184-634-402	Sequence 402, App
34	47.4	1.5	957	10	US-09-373-658-73	Sequence 73, Appl
35	47.4	1.5	957	11	US-09-989-687-73	Sequence 73, Appl
36	47.4	1.5	1089	15	US-10-267-430-1	Sequence 1, Appli
37	47.4	1.5	1137	15	US-10-184-644-518	Sequence 518, App
38	47.4	1.5	1137	15	US-10-184-634-518	Sequence 518, App
39	47.2	1.5	1523	15	US-10-184-644-290	Sequence 290, App
40	47.2	1.5	1523	15	US-10-184-634-290	Sequence 290, App
41	46.6	1.5	5001	15	US-10-156-761-2323	Sequence 2323, Ap
42	46.2	1.4	507	15	US-10-029-386-10811	Sequence 10811, A
43	46.2	1.4	552	15	US-10-029-386-5873	Sequence 5873, Ap
44	46.2	1.4	3336	15	US-10-156-761-376	Sequence 376, App
45	46.2	1.4	4309	13	US-10-170-385-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-10-024-368-1
; Sequence 1, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3202
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-024-368-1

Query Match	100.0%;	Score 3202;	DB 15;	Length 3202;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3202;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTTTACTACAGGATCCGAGCATTTCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA	60	
Db	1	TTTTACTACAGGATCCGAGCATTTCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA	60	
QY	61	CCTGGGTTGTTTGGCTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGGAACGC	120	
Db	61	CCTGGGTTGTTTGGCTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGGAACGC	120	
QY	121	CCTTCACTGCACAGAGGGATGGAGATGGAGTGGGAGTGGCCGGCAGCAGATCCTTGCCCG	180	
Db	121	CCTTCACTGCACAGAGGGATGGAGATGGAGTGGGAGTGGCCGGCAGCAGATCCTTGCCCG	180	

Db 361 TGCTGTTTCATCCTACCCACCCTAAAGGTGGGATCTTGGCCCTTGTGGGAAGTGCCAG 420

QY 421 GAGGGCCTGGAGGGGGTGCCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGCC 480

Db 421 GAGGGCCTGGAGGGGGTGCCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGCC 480

QY 481 TCTGGCCCCAGGGCCTGTCCCCCAGCCACCACACCAAGCTGAAGAAGACATGGCTACA 540

Db 481 TCTGGCCCCAGGGCCTGTCCCCCAGCCACCACCAAGCTGAAGAAGACATGGCTACA 540

QY 541 CGGCACCTCGGAGCAGTTTGAATGTCCACGGCTGCCCTGAGTFCGAGGAGGCCGGTT 600

Db 541 CGGCACCTCGGAGCAGTTTGAATGTCCACGGCTGCCCTGAGTFCGAGGAGGCCGGTT 600

QY 601 GTCGGCTCCGGGCCCTCAAAAGGCAGGCAGCCCCGAGGTCCAGGGAGCAATGGGCAGT 660

Db 601 GTCGGCTCCGGGCCCTCAAAAGGCAGGCAGCCCCGAGGTCCAGGGAGCAATGGGCAGT 660

QY 661 CCAGCCCCAAGCGGCCACCCGACCCCTTCCAGGCACTGCAGAACAGGGGCTGGGGST 720

Db 661 CCAGCCCCAAGCGGCCACCCGACCCCTTCCAGGCACTGCAGAACAGGGGCTGGGGST 720

QY 721 TGGCAGGAGGTTTCGGGACACATCGATAGGGAACAAGGATGTGACTCGGGACAGCATGAT 780

Db 721 TGGCAGGAGGTTTCGGGACACATCGATAGGGAACAAGGATGTGACTCGGGACAGCATGAT 780

QY 781 GAGCAGAAAGGACCCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 840

Db 781 GAGCAGAAAGGACCCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 840

QY 841 CCAATGCCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCAGCAGCTGGA 900

Db 841 CCAATGCCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCAGCAGCTGGA 900

QY 901 GAGGGAGGAGGGCACGCCTGCCACTCTCAGCAAGTGGAGATCGCCTCTGGGAGGGGAG 960

Db 901 GAGGGAGGAGGGCACGCCTGCCACTCTCAGCAAGTGGAGATCGCCTCTGGGAGGGGAG 960

QY 961 CTGCAGCAGGAGGAAGACACAGCCCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1020

Db 961 CTGCAGCAGGAGGAAGACACAGCCCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1020

QY 1021 CCTGACAGCCGGCTCAGCACAGGCCTGCCAAGCACCTGTCAGTGCTTGGGGGACCGA 1080

Db 1021 CCTGACAGCCGGCTCAGCACAGGCCTGCCAAGCACCTGTCAGTGCTTGGGGGACCGA 1080

QY 1081 CTGTGCCGCCTGCTGCGGAGGGAGCGGAGGCCCTGGCTGGGCCCCAGCGGAAGGCCAA 1140

Db 1081 CTGTGCCGCCTGCTGCGGAGGGAGCGGAGGCCCTGGCTGGGCCCCAGCGGAAGGCCAA 1140

QY 1141 GGGCCAGCCGTGACAGGGGACAGCCCAAGGCAATCCACGTGCTGCAGCCGTTGCCACCAT 1200

Db 1141 GGGCCAGCCGTGACAGGGGACAGCCCAAGGCAATCCACGTGCTGCAGCCGTTGCCACCAT 1200

QY 1201 GGACTCTTCAACACCCACTGGCGATGTCCCCGTGCAGCACCGGCTGTGTGGCCTGT 1260

Db 1201 GGACTCTTCAACACCCACTGGCGATGTCCCCGTGCAGCACCGGCTGTGTGGCCTGT 1260

QY 1261 GGTCTGTGGCAGGCACTGGSGGGCCAGGGAGAAAGCAGGCTTTCAGGAGCAGTCCGG 1320

Db 1261 GGTCTGTGGCAGGCACTGGSGGGCCAGGGAGAAAGCAGGCTTTCAGGAGCAGTCCGG 1320

QY 1321 GAGGAGTGCACGCAGGAGGCCGGGACGCTGCCTGTTCCCTGATGCTGACCCAGTTTGT 1380

Db 1321 GAGGAGTGCACGCAGGAGGCCGGGACGCTGCCTGTTCCCTGATGCTGACCCAGTTTGT 1380

QY 1381 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACAGGTCTGGGTCAAGTTTGAT 1440

Db 1381 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACAGGTCTGGGTCAAGTTTGAT 1440

QY 1441 ATCCGGGGCAGCTGCCCTGCCAAGCTGATGCCCGGTATGGCCCCCGGGATGCAGGC 1500

Db 1441 ATCCGGGGCAGCTGCCCTGCCAAGCTGATGCCCGGTATGGCCCCCGGGATGCAGGC 1500

QY 1501 CAGCAGAAAGGAATCAACACAGAAAAACGCCCCAACTCCACAACTTCTGCAATGGCGAC 1560

Db 1501 CAGCAGAAAGGAATCAACACAGAAAAACGCCCCAACTCCACAACTTCTGCAATGGCGAC 1560

QY 1561 ACCCAGAGCAACCAAGAGCATCAAAGAGGAGACCCCCGATTCGCTGAGACCCAGCAGAG 1620

Db 1561 ACCCAGAGCAACCAAGAGCATCAAAGAGGAGACCCCCGATTCGCTGAGACCCAGCAGAG 1620

QY 1621 GACCGTGTGGCCGAGGGCCCTGCTTGTCTTCTCTGCGAACTGTGGTCTTCTACC 1680

Db 1621 GACCGTGTGGCCGAGGGCCCTGCTTGTCTTCTCTGCGAACTGTGGTCTTCTACC 1680

QY 1681 GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACATGGGCTTCGCCCCCGTCACTCCG 1740

Db 1681 GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACATGGGCTTCGCCCCCGTCACTCCG 1740

QY 1741 GCGGTGCCCCAGTGATGACCGCATCACCAACATCCTGGACAGCATTATCGCACAGGTGGTG 1800

Db 1741 GCGGTGCCCCAGTGATGACCGCATCACCAACATCCTGGACAGCATTATCGCACAGGTGGTG 1800

QY 1801 GAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGCTTCGAGCTGGCCCGGTCTGCGC 1860

Db 1801 GAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGCTTCGAGCTGGCCCGGTCTGCGC 1860

QY 1861 AAGGGCCTGGCCCTCTCTCCAGTGGGGCCCCGGCTGCCCTCCCCAGGGGCTTTG 1920

Db 1861 AAGGGCCTGGCCCTCTCTCCAGTGGGGCCCCGGCTGCCCTCCCCAGGGGCTTTG 1920

QY 1921 CTGTGGCTGCAGGAGCCCCAGCCTTGCCCTCGGGCTGGCTTCCACCTCTTCCAGGAGCAC 1980

Db 1921 CTGTGGCTGCAGGAGCCCCAGCCTTGCCCTCGGGCTGGCTTCCACCTCTTCCAGGAGCAC 1980

QY 1981 TGGAGGCAGGGCCAGCCTGTGTGTGTCAGGGATCCAAAGGACATTCAGGGCAACTG 2040

Db 1981 TGGAGGCAGGGCCAGCCTGTGTGTGTCAGGGATCCAAAGGACATTCAGGGCAACTG 2040

QY 2041 TGGGGACAGAAAGCTCTTGGGGCACTTGGAGGCCAGGTGCAGGCGCTGAGCCCCCTCGGA 2100

Db 2041 TGGGGACAGAAAGCTCTTGGGGCACTTGGAGGCCAGGTGCAGGCGCTGAGCCCCCTCGGA 2100

QY 2101 CCTCCCCAGCCAGCAGCCTGGGCAGCACAACTTCTGGGAGGGCTTCTCCTGGCCTGAG 2160

Db 2101 CCTCCCCAGCCAGCAGCCTGGGCAGCACAACTTCTGGGAGGGCTTCTCCTGGCCTGAG 2160

QY 2161 CTTTGGCCCCAAAGTCAAGAGGGCTCTGTCTCTCTGCTGCACCGAGCTTGGGGGATGAG 2220

Db 2161 CTTTGGCCCCAAAGTCAAGAGGGCTCTGTCTCTCTGCTGCACCGAGCTTGGGGGATGAG 2220

QY 2221 GACACAGCAGGGTGGAGAACCTAGTGCCAGTCTGCGACTTCCGAGTACTCGGCCCTC 2280

Db 2221 GACACAGCAGGGTGGAGAACCTAGTGCCAGTCTGCGACTTCCGAGTACTCGGCCCTC 2280

QY 2281 CATGGAAAACTCAACCTGGCTTCTACCTCCCACCGGGCTTGGCCCTGCGTCCACTGGAG 2340

Db 2281 CATGGAAAACTCAACCTGGCTTCTACCTCCCACCGGGCTTGGCCCTGCGTCCACTGGAG 2340

QY 2341 CCCCAGCTCTGGCAGCCTATGGTGTGAGCCCCGACCCGGGACACCTGGGGACCAAGAAC 2400

Db 2341 CCCCAGCTCTGGCAGCCTATGGTGTGAGCCCCGACCCGGGACACCTGGGGACCAAGAAC 2400

QY 2401 CTCTGTGTGGAGGTGCCGACCTGGTCAAGCATCTTGGTGCATGCCGACACACCACTGCCT 2460

Db 2401 CTCTGTGTGGAGGTGCCGACCTGGTCAAGCATCTTGGTGCATGCCGACACACCACTGCCT 2460

QY 2461 GCCTGGCACCGGGCAGAAAGACTTCTCTTTCAGGCCCTGGACGGGAGGGCTCTGGTCT 2520

Db 2461 GCCTGGCACCGGGCAGAAAGACTTCTCTTTCAGGCCCTGGACGGGAGGGCTCTGGTCT 2520

QY 2521 CCGGGCAGCCAGGTGAGCACTGTGTGGCACGTGTTCGGGGCACAGGACGCCAGCGCATC 2580

Db 2521 CCGGGCAGCCAGGTGAGCACTGTGTGGCACGTGTTCGGGGCACAGGACGCCAGCGCATC 2580

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 13:21:41 ; Search time 216 Seconds
(without alignments)
8226.633 Million cell updates/sec

Title: US-10-024-368-1
Perfect score: 3202
Sequence: 1 ttttactacaagatccgag.....tgtgcccttttatgtacagg 3202

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3202	100.0	3202	4	US-09-287-354-1
2	52	1.6	1402	1	US-08-447-965A-1
3	50.6	1.6	1257	4	US-09-252-991A-1027
4	49.8	1.6	1692	4	US-09-252-991A-978
5	48.2	1.5	4411529	3	US-09-103-840A-1
6	47.6	1.5	7218	1	US-08-232-463-14
7	46	1.4	498	4	US-09-252-991A-1030
8	45.6	1.4	80161	3	US-09-036-987A-1
9	45.6	1.4	80161	3	US-09-370-700-1
10	45.6	1.4	430	4	US-09-603-207-1
11	44.8	1.4	430	4	US-09-621-976-16656
12	44.4	1.4	1512	4	US-09-252-991A-14111
13	44.4	1.4	1866	4	US-09-252-991A-14118
14	44.4	1.4	4287	4	US-09-252-991A-14160
15	43.8	1.4	30001	1	US-08-125-468-1
16	43.8	1.4	30001	2	US-08-474-933-1
17	43.6	1.4	5467	1	US-07-745-206A-12
18	43.6	1.4	5467	2	US-08-311-363-12
19	43.6	1.4	5640	4	US-09-620-312D-41
20	43.4	1.4	1254	4	US-09-252-991A-7536
21	43.4	1.4	1989	4	US-09-252-991A-7805
22	43.4	1.4	2109	4	US-09-370-838-153
23	43.4	1.4	2313	4	US-09-252-991A-7945
24	42.8	1.3	367	4	US-09-072-596-249
25	42.8	1.3	367	4	US-09-072-967-254
26	42	1.3	540	4	US-09-252-991A-10333
27	42	1.3	891	4	US-09-252-991A-10527

28	42	1.3	1320	4	US-09-252-991A-10535	Sequence 10535, A
29	42	1.3	1473	4	US-09-252-991A-10639	Sequence 10639, A
30	42	1.3	1723	4	US-09-614-034-187	Sequence 187, App
31	42	1.3	1723	4	US-09-016-434-1182	Sequence 1182, Ap
32	42	1.3	2055	4	US-09-252-991A-10433	Sequence 10433, A
33	42	1.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
34	42	1.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
35	41.8	1.3	364	4	US-09-621-976-17202	Sequence 17202, A
36	41.8	1.3	2210	3	US-08-464-700-1	Sequence 1, Appli
37	41.6	1.3	418	4	US-09-280-116-168	Sequence 168, App
38	41.4	1.3	3524	4	US-09-077-940A-3	Sequence 3, Appli
39	41.2	1.3	510	4	US-09-252-991A-5973	Sequence 5973, Ap
40	41.2	1.3	678	4	US-09-252-991A-10534	Sequence 10534, A
41	41.2	1.3	960	4	US-09-252-991A-15798	Sequence 15798, A
42	41.2	1.3	1506	4	US-09-252-991A-15926	Sequence 15926, A
43	41.2	1.3	1596	4	US-09-252-991A-5938	Sequence 5938, Ap
44	41.2	1.3	1674	4	US-09-252-991A-5894	Sequence 5894, Ap
45	41.2	1.3	1890	4	US-09-252-991A-10529	Sequence 10529, A

ALIGNMENTS

RESULT 1
US-09-287-354-1
; Sequence 1, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3202
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-287-354-1

Query Match	100.0%;	Score 3202;	DB 4;	Length 3202;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3202;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Qy	1	TTTTACTAAGGATCCGAGCATCCAGGTTGGCAAGGAGCCCTTGGCAGCTGCGGAA	60	
Db	1	TTTTACTAAGGATCCGAGCATCCAGGTTGGCAAGGAGCCCTTGGCAGCTGCGGAA	60	
Qy	61	CCTGGGTTGTTGGCTTAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGAAACGC	120	
Db	61	CCTGGGTTGTTGGCTTAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGAAACGC	120	
Qy	121	CCTTCACTGCACGAGGGATGGAGAGATGGGAGCTGGCCGGCAGCAGAATCCTTGCCCG	180	
Db	121	CCTTCACTGCACGAGGGATGGAGAGATGGGAGCTGGCCGGCAGCAGAATCCTTGCCCG	180	
Qy	181	CTCTTCTGGGGCAGCCAGACACTGTGCTGCTGGACCTCTGCGCCGCTTGCCCCCAGGC	240	
Db	181	CTCTTCTGGGGCAGCCAGACACTGTGCTGCTGGACCTCTGCGCCGCTTGCCCCCAGGC	240	
Qy	241	CTTGTTCACTCTTTGGCAACGCTCTGGGCTGGGCCAGCGATGGGAACCTTGGGTACCAG	300	
Db	241	CTTGTTCACTCTTTGGCAACGCTCTGGGCTGGGCCAGCGATGGGAACCTTGGGTACCAG	300	
Qy	301	CTGGGGCCACCAGCAACACCAAGGTGCCCTCTCCTGAGCCGCTCTCACCAGCGGGGC	360	
Db	301	CTGGGGCCACCAGCAACACCAAGGTGCCCTCTCCTGAGCCGCTCTCACCAGCGGGGC	360	
Qy	361	TGCTGTTCATCCTTACCCACCCACTAAAGGTGGGATCTTGGCCCTTGTGGGAAGTCCAG	420	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2004, 13:08:50 ; Search time 7850 Seconds
(without alignments)
12180.726 Million cell updates/sec

Title: US-10-024-368-1
Perfect score: 3202
Sequence: 1 ttttactacaagatccgag.....tgtgcccttttatgtacagg 3202

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vrt:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_phg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1273	39.8	2969	11	AK076280 Mus muscu
2	781.4	24.4	1094	12	BM799906 AGENCOURT
3	764.8	23.9	933	13	BQ961259 AGENCOURT
4	736.2	23.0	811	12	BG739858 602630590

5	586.2	21.4	1062	12	BI197548
6	665.8	20.8	837	13	BX341278
7	659.4	20.6	869	12	BG743219
8	657	20.5	989	10	BF316119
9	636.2	19.9	780	12	BI827477
10	628.8	19.6	725	12	BG677197
11	621.2	19.4	858	12	BG330970
12	609	19.0	978	10	BF346259
13	594.4	18.6	951	12	BI261186
14	574	17.9	586	14	CB154516
15	566	17.7	802	10	BE276310
16	556.2	17.4	1068	12	BG676990
17	554	17.3	861	10	BE784838
18	549.6	17.2	618	10	BE410741
19	523.2	16.3	867	13	BQ951435
20	520	16.2	775	12	BG923878
21	481.2	15.0	855	13	BX415340
22	440	13.7	654	12	BG963522
23	439.6	13.7	921	13	BQ925145
24	428.4	13.4	783	13	BQ444010
25	424.4	13.3	554	10	BF077353
26	424.2	13.2	732	12	BI655486
27	418.6	13.1	875	12	BI687595
28	418.6	13.1	909	10	BF528824
29	394	12.3	700	13	BQ443965
30	367.2	11.5	623	10	BB612436
31	360.4	11.3	515	14	CD548907
32	353.4	11.0	789	10	BE274754
33	349.8	10.9	894	13	BQ217662
34	336.8	10.5	624	10	BE573367
35	330.8	10.3	471	9	AA760207
36	321	10.0	481	13	BQ307325
37	313.2	9.8	490	14	R67180
38	308.8	9.6	771	12	BI665435
39	307.6	9.6	543	14	CB612355
40	306	9.6	916	13	BUI96087
41	301.6	9.4	867	12	BI655258
42	285	8.9	389	10	BF225274
43	283.4	8.9	820	28	AQ781018
44	268.8	8.4	474	12	BG992856
45	266.8	8.3	311	10	AW354018

ALIGNMENTS

RESULT 1	AK076280	AK076280	2969 bp	mRNA	linear	HTC 18-SEP-2003
LOCUS	Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632409G21 product:hairless, full insert sequence.					
DEFINITION	AK076280 GI:26096706					
ACCESSION	AK076280.1					
VERSION	HTC; CAP trapper.					
KEYWORDS	Mus musculus (house mouse)					
SOURCE	Mus musculus					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
REFERENCE	1					
AUTHORS	Carninci,P. and Hayashizaki,Y.					
TITLE	High-efficiency full-length cDNA cloning					
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)					
MEDLINE	99279253					
PUBMED	10349636					
REFERENCE	2					
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.					
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes					
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)					
MEDLINE	20499374					
PUBMED	11042159					
REFERENCE	3					

Db 361 CCCAGGCATTCCACGCTGCTGTCAGCCGTTGCCACCATGGACTCTTCAACACCCACTGGCG 420
QY 1224 ATGTCCCGCTGCAGCCACCGGCTGTGTGTGGCTGTGGTCTGTGGTGGCAGGACTGGCG 1283
Db 421 ATGTCCCGCTGCAGCCACCGGCTGTGTGTGGCTGTGGTCTGTGGTGGCAGGACTGGCG 480
QY 1284 GGCCAGGGAGAAAGCAGGCTTTTCAGGAGCAGTCCGCGGAGGAGTGTCACGAGGAGGCC 1341
Db 481 GGCCAGGGAGAAAGCAGGCTTTTCAGGAGCAGTCCGCGGAGGAGTGTCACGAGGAGGCC 538

RESULT 7
ABS69028/c
ID ABS69028 standard; DNA; 335 BP.
XX
AC ABS69028;
XX
DT 21-NOV-2002 (first entry)
XX
DE Novel murine polynucleotide isolated using gene trap technology #91.
XX
KW Mouse; gene trapped sequence; GTS; functional genomic analysis;
KW phage display system; gene chip; temporal gene expression;
KW tissue specific gene expression; antisense inhibition; gene targeting;
KW development disorder; cell differentiation disorder; aging; cancer;
KW autoimmune disease; lupus; inflammatory disorder; skin disorder;
KW degenerative disorder; ds.
XX
OS Mus musculus.
XX
PN US2002102543-A1.
XX
PD 01-AUG-2002.
XX
PF 30-NOV-2000; 2000US-00728445.
XX
PR 01-DEC-1999; 99US-0168358P.
XX
PA (FRIE/) FRIEDRICH G.
PA (ZAMB/) ZAMBROWICZ B.
PA (SAND/) SANDS A T.
XX
PI Friedrich G, Zambrowicz B, Sands AT;
XX
DR WPI; 2002-690598/74.
XX
PT Novel murine polynucleotides that individually identify novel genes into
PT which a retroviral gene trap vector has integrated, useful in genomic
PT analysis and in discovery, development of therapeutic and diagnostic
PT agents.
XX
PS Claim 1; Page 52; 296pp; English.
XX
CC The invention describes an isolated murine polynucleotide (I) comprising
CC a contiguous stretch of at least 60 nucleotides of one of 265-677
CC nucleotide 891 OMNIBANK gene trapped sequences (GTSs) (S), given in the
CC specification. The novel genes and cells are useful in functional genomic
CC analysis and in the discovery and development of new therapeutic and
CC diagnostic agents and methods. (I) is useful for identifying the coding
CC regions of the murine genome, to isolate cDNAs, genomic clones, or full-
CC length genes/polynucleotides or homologues, heterologues, paralogues, or
CC orthologues that are capable of hybridising to one or more of the GTSs
CC under stringent conditions. (I) can be incorporated into a phage display
CC system that can be used to screen for proteins, or other ligands, that
CC are capable of binding an amino acid sequence encoded by an
CC oligonucleotide or polynucleotide sequence in at least one of the TS
CC sequences. (I) is useful in addressable arrays, such as gene chips, to
CC identify and characterise temporal and tissue specific gene expression,
CC to identify the gene of interest from many sources and for genetic
CC manipulations such as antisense inhibition and gene targeting. Decreasing
CC the level of expression of (I) and/or down regulating the activity of
CC peptides or proteins encoded by (I) is useful for treating development
CC and cell differentiation disorders, aging, cancer, autoimmune disease,

CC lupus, inflammatory disorders, skin disorders and degenerative disorders.
CC This sequence represents a murine cDNA isolated using gene trap
CC technology
XX
SQ Sequence 335 BP; 84 A; 77 C; 84 G; 74 T; 0 U; 16 Other;
Query Match 2.7%; Score 86.8; DB 6; Length 335;
Best Local Similarity 82.0%; Pred. No. 4.2e-08;
Matches 105; Conservative 0; Mismatches 22; Indels 1; Gaps 1;
QY 1134 AGGCCAAGGGCCAGCCGTCAGACGGGGACAGCCAGGCATTCACGCTGCTGCAGCCGTTG 1193
Db 128 AGGCCANGGGCCAGCCATGNCAGAGGACAGCCAGGCATTCNCATTGCNGCANNAGATG 69
QY 1194 CCACCATGGACTCTTCAACACCCACTGGCGATGTCCCGC-TGCAGGCCACCGGCTGTGTG 1252
Db 68 CCACCACGACTCTTCAACANCCNCTGNAGATGTTNCCACATGTAGCCACCGGCTGTGTG 9
QY 1253 TGGCCTGT 1260
Db 8 TAGCCTGT 1
RESULT 8
ABX41049
ID ABX41049 standard; cDNA; 429 BP.
XX
AC ABX41049;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #6214.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-00960352.
XX
PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
PS Claim 2; SEQ ID NO 6214; 245pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or

QY 61 CCTGGGTTGTTGGCTTAAACTCTGGTGGGCACTGTGAGAGAGCCGGGAGGCCGAACGC 120
Db |||||
676 CCTGGGTTGTTGGCTTAAACTCTGGTGGGCACTGTGAGAGAGCCGGGAGGCCGAACGC 735
QY 121 CCTTCACTGACACAGAGGGATGGAGAGATGGAGACTGGCCGGCAGCAGAAATCCTTGCCCG 180
Db |||||
736 CCTTCACTGACACAGAGGGATGGAGAGATGGAGACTGGCCGGCAGCAGAAATCCTTGCCCG 795
QY 181 CTCCTTCCCTGGGCGAGCCAGACACTGTGCTCCCTGGACCTCTCTGGCCCGCTTGTCCCCCAGGC 240
Db |||||
796 CTCCTTCCCTGGGCGAGCCAGACACTGTGCTCCCTGGACCTCTCTGGCCCGCTTGTCCCCCAGGC 855
QY 241 CTTGTTTCATACTCTTTGGCAACGCTCTGGCTGGGCGCAGGCGATGGGAACCTTGGGTACCA 300
Db |||||
856 CTTGTTTCATACTCTTTGGCAACGCTCTGGCTGGGCGCAGGCGATGGGAACCTTGGGTACCA 915
QY 301 CTGGGSCACACAGCAACACCAAGGTGCCCTCTCTGAGCCGCTGTCAACCCAGCGGGGC 360
Db |||||
916 CTGGGSCACACAGCAACACCAAGGTGCCCTCTCTGAGCCGCTGTCAACCCAGCGGGGC 975
QY 361 TGCTGTTTCATCCTACCCACCCACTAAAGGTGGGATCTTGGCCCTTGTGGGAAGTGCCAG 420
Db |||||
976 TGCTGTTTCATCCTACCCACCCACTAAAGGTGGGATCTTGGCCCTTGTGGGAAGTGCCAG 1035
QY 421 GAGGCTTGAGGGGGTGCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAAACAGGCC 480
Db |||||
1036 GAGGCTTGAGGGGGTGCAGTGGAGCCAGCGAACCCAGCGAGGAAGTGAAACAGGCC 1095
QY 481 TCTGGCCCCAGGGCCTGTCCCCCAGCCACACCAAGCTGAAGAAGACATGGCTCACA 540
Db |||||
1096 TCTGGCCCCAGGGCCTGTCCCCCAGCCACACCAAGCTGAAGAAGACATGGCTCACA 1155
QY 541 CGGCACTCGGAGCAGTTTGAATGTCCACGGGCTGCCCTGAGTTCAGGAGAGGCCGGTT 600
Db |||||
1156 CGGCACTCGGAGCAGTTTGAATGTCCACGGGCTGCCCTGAGTTCAGGAGAGGCCGGTT 1215
QY 601 GCTCGGCTCCGGGCCCTCAAAAGGCGAGGCAGCCCCGAGGTCCAGGGAGCAATGGGCAGT 660
Db |||||
1216 GCTCGGCTCCGGGCCCTCAAAAGGCGAGGCAGCCCCGAGGTCCAGGGAGCAATGGGCAGT 1275
QY 661 CCAGCCCCAAGCGGCCACCGGACCTTTCCAGGCACTGACAGCACTGGGGCTGGGGGT 720
Db |||||
1276 CCAGCCCCAAGCGGCCACCGGACCTTTCCAGGCACTGACAGCACTGGGGGTGGGGGT 1335
QY 721 TGGCAGGAGGTTCCGGACACATCGATAGGGAACAAGGATGTGACTCGGGACAGCATGAT 780
Db |||||
1336 TTGAGGAGGTCGGGACACATCGATAGGGAACAAGGATGTGACTCGGGACAGCATGAT 1395
QY 781 GAGCAGAAAGGACCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGACTTCAGGACATA 840
Db |||||
1396 GAGCAGAAAGGACCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGACTTCAGGACATA 1455
QY 841 CCATGCTCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCCAGGCAGCTGGA 900
Db |||||
1456 CCATGCTCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCCAGGCAGCTGGA 1515
QY 901 GAGGAGGAGGGCAGCCCTGCCACTCTCAGCAAGTGGGAGATGCGCTCTGGGAGGGAG 960
Db |||||
1516 GAGGAGGAGGGCAGCCCTGCCACTCTCAGCAAGTGGGAGATGCGCTCTGGGAGGGAG 1575
QY 961 CTGACAGGAGGAAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1020
Db |||||
1576 CTGACAGGAGGAAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1635
QY 1021 CTGACAGCCGGCTCAGCACAGGCTCGCCAAGCACCTGCTCAGTGGTTTGGGGACCGA 1080
Db |||||
1636 CTGACAGCCGGCTCAGCACAGGCTCGCCAAGCACCTGCTCAGTGGTTTGGGGACCGA 1695
QY 1081 CTGTGCCGCTGCTGCGGAGGAGCGGAGGCCCTGGCTTGGGCCCAGCGGGAAGGCCAA 1140
Db |||||
1696 CTGTGCCGCTGCTGCGGAGGAGCGGAGGCCCTGGCTTGGGCCCAGCGGGAAGGCCAA 1755

QY 1141 GGGCCAGCCGTGACAGGGGACAGCCAGGCATTCACGCTGCTGCAGCCGTTGCCACCAT 1200
Db |||||
1756 GGGCCAGCCGTGACAGAGGACAGCCAGGCATTCACGCTGCTGCAGCCGTTGCCACCAT 1815
QY 1201 GGACTCTTCAACACCCACTGGCGATGTCCCCTGTCAGCCACCGGCTGTGTGGCCTGT 1260
Db |||||
1816 GGACTCTTCAACACCCACTGGCGATGTCCCCTGTCAGCCACCGGCTGTGTGGCCTGT 1875
QY 1261 GGTCTGTGGCAGGCACCTGGCGGSCCAGGAGAAAGCAGGCTTTCAGGAGCAGTCCGCG 1320
Db |||||
1876 GGTCTGTGGCAGGCACCTGGCGGSCCAGGAGAAAGCAGGCTTTCAGGAGCAGTCCGCG 1935
QY 1321 GAGGAGTGCACGCAGGAGGCCCGGCACGCTGCCTGTTCCCTGATGCTGACCCAGTTTGT 1380
Db |||||
1936 GAGGAGTGCACGCAGGAGGCCCGGCACGCTGCCTGTTCCCTGATGCTGACCCAGTTTGT 1995
QY 1381 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACACAGTCTGGGTCAAGTTTGT 1440
Db |||||
1996 TCCAGCCAGGCTTTGGCAGAGCTGAGCACTGCAATGCACACAGTCTGGGTCAAGTTTGT 2055
QY 1441 ATCCGGGGCACTGCCCTGCCAAGCTGATGCCCGGTATGGGCCCCCGGGGATGCAGGC 1500
Db |||||
2056 ATCCGGGGCACTGCCCTGCCAAGCTGATGCCCGGTATGGGCCCCCGGGGATGCAGGC 2115
QY 1501 CAGCAGAAAGAAATCAACACAGAAAACGCCCCCAACTCCACAACCTTCCTGCAATGGCGAC 1560
Db |||||
2116 CAGCAGAAAGAAATCAACACAGAAAACGCCCCCAACTCCACAACCTTCCTGCAATGGCGAC 2175
QY 1561 ACCCAGAGACCAAGAGCATCAAAAGAGGAGACCCCGGATTCGGCTGAGACCCCGAGAG 1620
Db |||||
2176 ACCCAGAGACCAAGAGCATCAAAAGAGGAGACCCCGGATTCGGCTGAGACCCCGAGAG 2235
QY 1621 GACCGTCTGGCCGAGGGCCCTTGTCTCTCTCTGCGAACTGCTGGCTTCTTACC 1680
Db |||||
2236 GACCGTCTGGCCGAGGGCCCTTGTCTCTCTCTGCGAACTGCTGGCTTCTTACC 2295
QY 1681 GCGGTCAAACTCTGCTTGGGCCATGAGCGAATACATATGGCTTCGCCCCCGTCACTCCG 1740
Db |||||
2296 GCGGTCAAACTCTGCTTGGGCCATGAGCGGATACATATGGCTTCGCCCCCGTCACTCCG 2355
QY 1741 GCCCTGCCCAGTGATGACCCGATCACCAACATCCTGGACAGCATATCGCACAGGTGGTG 1800
Db |||||
2356 GCCCTGCCCAGTGATGACCCGATCACCAACATCCTGGACAGCATATCGCACAGGTGGTG 2415
QY 1801 GAAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGCTTCAGCTGGCCCGGCTCGCGC 1860
Db |||||
2416 GAAACGGAAGATCCAGGAGAAAGCCCTGGGGCCGGGCTTCAGCTGGCCCGGCTCGCGC 2475
QY 1861 AAGGCTTGGGCTGCCCTCTCTCCAGTGGGCCCCCGGCTGCCCTCCCCCAGGGGCTTG 1920
Db |||||
2476 AAGGCTTGGGCTGCCCTCTCTCCAGTGGGCCCCCGGCTGCCCTCCCCCAGGGGCTTG 2535
QY 1921 CTGTGGCTGCAGGAGCCCCAGCCTTGCCCTCGGGCGTGGCTTCCACCTCTTCCAGGAGCAC 1980
Db |||||
2536 CTGTGGCTGCAGGAGCCCCAGCCTTGCCCTCGGGCGTGGCTTCCACCTCTTCCAGGAGCAC 2595
QY 1981 TGGAGGCAGGGCCAGCCTGTGTGGTGTCAAGGATCCAAAGGACATTCAGGGCAACCTG 2040
Db |||||
2596 TGGAGGCAGGGCCAGCCTGTGTGGTGTCAAGGATCCAAAGGACATTCAGGGCAACCTG 2655
QY 2041 TGGGGGACAGAAGCTTGGGGCACTTGGAGGCAGGTGAGGCGCTGAGCCCTCGGA 2100
Db |||||
2656 TGGGGGACAGAAGCTTGGGGCACTTGGAGGCAGGTGAGGCGCTGAGCCCTCGGA 2715
QY 2101 CCTCCCCAGCCAGCAGCCTGGGAGCAACAACATTTGGGAGGGCTTCTCTGGCCTGAG 2160
Db |||||
2716 CCTCCCCAGCCAGCAGCCTGGGAGCAACAACATTTGGGAGGGCTTCTCTGGCCTGAG 2775
QY 2161 CTTGCCCCAAAGTCAAGAGGGCTCTGTCTCTCTGCTGCAACCGAGCTTTGGGGGATGAG 2220
Db |||||
2776 CTTGCCCCAAAGTCAAGAGGGCTCTGTCTCTCTGCTGCAACCGAGCTTTGGGGGATGAG 2835
QY 2221 GACACCAGCAGGGTGGAGAACCTAGTGCACGCTGCGCACTTCGGGAGTACTGCGGCCCTC 2280

QY 1981 TGGAGGCAGGGCCAGCCTGTCTTGGTGTCTAGGGATCCAAAGGACATTGCAGGSCAACCTG 2040
Db |||||
4080 TGGAGGCAGGGCCAGCCTGTCTTGGTGTCTAGGGATCCAAAGGACATTGCAGGSCAACCTG 4139
QY 2041 TGGGGGACAGAACTCTTGGGGCAGCTTGGAGGCGCAGGTGCAGGGCGCTGAGCCCTCGGA 2100
Db |||||
4140 TGGGGGACAGAACTCTTGGGGCAGCTTGGAGGCGCAGGTGCAGGGCGCTGAGCCCTCGGA 4199
QY 2101 CCTCCCCAGCCAGCAGCTGGGCAGCACAACTTCTGGGAGGGCTTCTCTGGCGCTGAG 2160
Db |||||
4200 CCTCCCCAGCCAGCAGCTGGGCAGCACAACTTCTGGGAGGGCTTCTCTGGCGCTGAG 4259
QY 2161 CTTGCCCCAAAGTCAGACGAGGGCTCTGTCTCTCTGCTGCACCGAGCTTTGGGGGATGAG 2220
Db |||||
4260 CTTGCCCCAAAGTCAGACGAGGGCTCTGTCTCTCTGCTGCACCGAGCTTTGGGGGATGAG 4319
QY 2221 GACACCAGCAGGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCTC 2280
Db |||||
4320 GACACCAGCAGGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCTC 4379
QY 2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCTTGCCTGCGTCCACTGGAG 2340
Db |||||
4380 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCTTGCCTGCGTCCACTGGAG 4439
QY 2341 CCCCAGCTCTGGGCAGCCTATGCTGTGAGCCCGCACCGGGGACACCTGGGGACCAAGAAC 2400
Db |||||
4440 CCCCAGCTCTGGGCAGCCTATGCTGTGAGCCCGCACCGGGGACACCTGGGGACCAAGAAC 4499
QY 2401 CTCTGTGTGGAGGTGGCCGACCTGTGTGATGATCTCTGTGTGATGCCGACACACACTGCCT 2460
Db |||||
4500 CTCTGTGTGGAGGTGGCCGACCTGTGTGATGATCTCTGTGTGATGCCGACACACACTGCCT 4559
QY 2461 GCCTGGCACCGGGCACAGAAAGACTTCTTTTTCAGGCTTGGACGGGAGGGGCTCTGGTCT 2520
Db |||||
4560 GCCTGGCACCGGGCACAGAAAGACTTCTTTTTCAGGCTTGGACGGGAGGGGCTCTGGTCT 4619
QY 2521 CCGGCAGCCAGGTGAGCACTGTGTGGCACGTGTTCGGGGCACAGGACGCCAGCGCATC 2580
Db |||||
4620 CCGGCAGCCAGGTGAGCACTGTGTGGCACGTGTTCGGGGCACAGGACGCCAGCGCATC 4679
QY 2581 CGCGCTTCTCCAGATGGTGTGCGGCGGGGCGGCGGCTGAGGCTTGGCGCCCCA 2640
Db |||||
4680 CGCGCTTCTCCAGATGGTGTGCGGCGGGGCGGCGGCTGAGGCTTGGCGCCCCA 4739
QY 2641 GGCAGCTGTACTGTGATGAGGGCTGCGGCGGCGGCTGCGGAGGAGTGGGGCGTGAGC 2700
Db |||||
4740 GGCAGCTGTACTGTGATGAGGGCTGCGGCGGCGGCTGCGGAGGAGTGGGGCGTGAGC 4799
QY 2701 TGCTGGACCTGTCTCAGGCCCCCGGAGAGGCGGCTGTGTGTGCTGAGGGGCTCCCCAC 2760
Db |||||
4800 TGCTGGACCTGTCTCAGGCCCCCGGAGAGGCGGCTGTGTGTGCTGAGGGGCTCCCCAC 4859
QY 2761 CAGGTGCAGGGCCTGGTGCAGCAGTGCAGCGTCACTCAGCACTTCTCTCTCCCTGAGACC 2820
Db |||||
4860 CAGGTGCAGGGCCTGGTGCAGCAGTGCAGCGTCACTCAGCACTTCTCTCTCCCTGAGACC 4919
QY 2821 TCTGCCCTCTGTCTCAGCTCTGCCACAGGACCCAGCCTTCCCCCTGACTGCCACCTG 2880
Db |||||
4920 TCTGCCCTCTGTCTCAGCTCTGCCACAGGACCCAGCCTTCCCCCTGACTGCCACCTG 4979
QY 2881 CTTTATGCCAGATGGACTGGGCTGTGTCCAGCAGTGAAGTGGCGGTGGGACATTA 2940
Db |||||
4980 CTTTATGCCAGATGGACTGGGCTGTGTCCAGCAGTGAAGTGGCGGTGGGACATTA 5039
QY 2941 CAGGAGGCCAAATAGAGGGATGCTAGGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG 3000
Db |||||
5040 CAGGAGGCCAAATAGAGGGATGCTAGGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG 5099
QY 3001 GTGCTCAGCCCAGGCACAACTTTCAGAGGGGATGGCGCTAGGGGACTTGGGGATTTCTGG 3060
Db |||||
5100 GTGCTCAGCCCAGGCACAACTTTCAGAGGGGATGGCGCTAGGGGACTTGGGGATTTCTGG 5159

QY 3061 TCAACCCCAAGCACCACCTCTTGGGCACAAGCAGGGCAGCTGTCTCCCTCCCTTTAAG 3120
Db |||||
5160 TCAACCCCAAGCACCACCTCTTGGGCACAAGCAGGGCAGCTGTCTCCCTCCCTTTAAG 5219
QY 3121 CCAACAACCAAGTGCACCAAGCTCACACCTGTCTCTCAGGCTGGCATCTCCCCAC 3180
Db |||||
5220 CCAACAACCAAGTGCACCAAGCTCACACCTGTCTCTCAGGCTGGCATCTCCCCAC 5279
QY 3181 CCTGTGCCCTTTT 3193
Db |||||
5280 CCTGTGCCCTTT 5292

RESULT 5
AAZ06349
ID AAZ06349 standard; DNA; 3567 BP.
XX AAZ06349;
AC AC
XX XX
DT 09-NOV-1999 (first entry)
TX Human Hairless wildtype protein nucleic acid sequence.
DE alopecia; congenital alopecia; congenital atrichia;
XX androgenetic alopecia; alopecia areata; alopecia universalis; wildtype;
KW hair follicle; ds.
KW Homo sapiens.
XX OS
XX WO9938965-A1.
PN 05-AUG-1999.
XX PD
XX 29-JAN-1999; 99WO-US002128.
XX PF
XX 29-JAN-1998; 98US-0073043P.
PR (UYCO) UNIV COLUMBIA NEW YORK.
XX Christiano AM;
XX WPI; 1999-479184/40.
DR P-PSDB; AAY15217.
XX Human hairless gene and protein, useful for identifying modulators of
PT hair growth.
PT Claim 7; Fig 6; 127pp; English.
PS This is the nucleotide sequence for the wildtype human hairless gene. The
XX gene was discovered by genotyping a Pakistani kindred (comprising of 4
CC affected males and 7 affected females) with an inherited form of
CC congenital alopecia universalis. The pedigree is strongly suggestive of
CC autosomal recessive inheritance. The invention provides methods and
CC sequences for the recombinant production of wild-type human hairless,
CC mutant human hairless and wild-type human whn (winged-helix-nude)
CC proteins, assays for screening for binding compounds, modulators and
CC homologues, and animal models of hairlessness. Human hairless conditions
CC such as androgenetic alopecia (male pattern baldness), alopecia areata,
CC alopecia totalis, congenital alopecia universalis, congenital alopecia
CC and severe T-cell immunodeficiency can be treated with compounds
CC identified in the assays. The methods are also useful for identifying
CC compounds that can be used to inhibit hair growth
XX
SQ Sequence 3567 BP; 656 A; 1173 C; 1148 G; 590 T; 0 U; 0 Other;

Query Match 91.7%; Score 2937.6; DB 2; Length 3567;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2943; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 TTTTACTACAGGATCCGAGCATTTCCAGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA 60
Db |||||
616 TTTTACTACAGGATCCGAGCATTTCCAGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA 675

CC	more lenient consensus cleavage site requirements and are less likely to
CC	degrade, in vivo, that hammerhead ribozymes. The catalytic nucleic acids
CC	are useful in pharmaceutical compositions for inhibiting hair production
CC	by a hair-producing cell, for inhibiting hair growth and for inhibiting
CC	the transition of a hair follicle from the anagen phase to the catagen
CC	phase. A non-human transgenic mammal is useful as a model for testing
CC	hair removal products which function by inhibiting hairless protein
CC	expression. The sequence presented is the human hairless protein cDNA
XX	
SQ	Sequence 5709 BP; 1015 A; 1950 C; 1826 G; 918 T; 0 U; 0 Other;
Query Match 99.5%; Score 3185; DB 7; Length 5709;	
Best Local Similarity 99.8%; Pred. No. 0;	
Matches 3188; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
QY	1 TTTTACTACAAGGATCCGAGCATTCCTCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA 60
Db	2100 TTTTACTACAAGGATCCGAGCATTCCTCCAGGTTGGCAAAGGAGCCCTTGGCAGCTGCGGAA 2159
QY	61 CCTGGGTGTTTGGCTTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGAACGC 120
Db	2160 CCTGGGTGTTTGGCTTAAACTCTGGTGGGCACCTGCAGAGAGCCGGGAGGCCGAACGC 2219
QY	121 CCTTCACTGCACAGAGGGATGGAGATGGGAGCTGGCCGGCAGCAGAAATCCTTGCCCG 180
Db	2220 CCTTCACTGCACAGAGGGATGGAGATGGGAGCTGGCCGGCAGCAGAAATCCTTGCCCG 2279
QY	181 CTCTTCCTGGGGCAGCCAGACACTGTGCCCTGGACCTCTTGGCCCGCTTGTCCCCAGGC 240
Db	2280 CTCTTCCTGGGGCAGCCAGACACTGTGCCCTGGACCTCTTGGCCCGCTTGTCCCCAGGC 2339
QY	241 CTGTGTCATACTCTTGGCAACGTCTGGGCTGGGCCAGGCGATGGAAACCTTGGGTACCAG 300
Db	2340 CTGTGTCATACTCTTGGCAACGTCTGGGCTGGGCCAGGCGATGGAAACCTTGGGTACCAG 2399
QY	301 CTGGGGCCACCAGCAACACCAAGGTGCCCTCTCTCTGAGCCCGCTGTCAACCCAGCGGGC 360
Db	2400 CTGGGGCCACCAGCAACACCAAGGTGCCCTCTCTCTGAGCCCGCTGTCAACCCAGCGGGC 2459
QY	361 TGCTGTTTCATCTACCCACCCACTAAAGGTGGGATCTTGGCCCTTGTGGGAAGTGCCAG 420
Db	2460 TGCTGTTTCATCTACCCACCCACTAAAGGTGGGATCTTGGCCCTTGTGGGAAGTGCCAG 2519
QY	421 GAGGGCTTGAGGGGGTGCCAGTGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGCC 480
Db	2520 GAGGGCTTGAGGGGGTGCCAGTGAGCCAGCGAACCCAGCGAGGAAGTGAACAAGGCC 2579
QY	481 TCTGGCCCCAGGGCCTGTCCCCCAGCCACACCAAGCTGAAGAAGACATGGCTCACA 540
Db	2580 TCTGGCCCCAGGGCCTGTCCCCCAGCCACACCAAGCTGAAGAAGACATGGCTCACA 2639
QY	541 CGGCACTCGGAGCAGTTTGAATGTCCACGGCTGCCCTGAGTTCGAGGAGAGGCCGGTT 600
Db	2640 CGGCACTCGGAGCAGTTTGAATGTCCACGGCTGCCCTGAGTTCGAGGAGAGGCCGGTT 2699
QY	601 GCTCGGCTCCGGGCCCTCAAAAGGCGAGGCAGGCCCGGAGGTCCAGGGAGCAATGGGCAGT 660
Db	2700 GCTCGGCTCCGGGCCCTCAAAAGGCGAGGCAGGCCCGGAGGTCCAGGGAGCAATGGGCAGT 2759
QY	661 CCAGCCCCAAGCGGCCACCGGACCTTTCCAGGCACTGCAGAACAGGGGCTGGGGGT 720
Db	2760 CCAGCCCCAAGCGGCCACCGGACCTTTCCAGGCACTGCAGAACAGGGGCTGGGGGT 2819
QY	721 TGGCAGGAGGTTCCGGACACATCGATAGGGAACAAGGATGTGGACTCGGGACAGCATGAT 780
Db	2820 TGGCAGGAGGTTCCGGGACACATCGATAGGGAACAAGGATGTGGACTCGGGACAGCATGAT 2879
QY	781 GAGCAGAAAGGACCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 840
Db	2880 GAGCAGAAAGGACCCCAAGATGGCCAGGCCAGTCTCCAGGACCCGGGACTTCAGGACATA 2939
QY	841 CCATGCTCTGCTTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCCAGGCAGCTGGA 900
Db	4020 CTGTGGCTGCAGGAGCCCGAGCCTTGGCCCTCGGCGTGGCTTCCACCTCTTCCAGGAGCAC 4079

Db	2940 CCATGCTGGCTCTCCCTGCAAAACTGGCTCAATGCCAAAGTTGTGCCCAGGCAGCTGGA 2999
QY	901 GAGGAGAGGGGCACGCTGCCACTCTCAGCAAGTGCAGAGATCGCTCTGGGAGGGGAG 960
Db	3000 GAGGAGAGGGGCACGCTGCCACTCTCAGCAAGTGCAGAGATCGCTCTGGGAGGGGAG 3059
QY	961 CTGACGAGGAGGAAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 1020
Db	3060 CTGACGAGGAGGAAGACACAGCCACCAACTCCAGCTCTGAGGAAGGCCAGGGTCCGGC 3119
QY	1021 CCTGACAGCCGGCTCAGCACAGGCTCGCAAGCACCTGCTCAGTGGTTTGGGGACCGA 1080
Db	3120 CCTGACAGCCGGCTCAGCACAGGCTCGCAAGCACCTGCTCAGTGGTTTGGGGACCGA 3179
QY	1081 CTGTGCCGCTGTGCGGAGGGAGCGGAGGCCCTTGGCTTGGGCCAGCGGAAGGCCAA 1140
Db	3180 CTGTGCCGCTGTGCGGAGGGAGCGGAGGCCCTTGGCTTGGGCCAGCGGAAGGCCAA 3239
QY	1141 GGGCCAGCCGTGACAGGGGACAGCCAGGCATTCCACGCTGTGACGCCGTTGCCACCAT 1200
Db	3240 GGGCCAGCCGTGACAGAGGACAGCCAGGCATTCCACGCTGTGACGCCGTTGCCACCAT 3299
QY	1201 GGACTCTTCAACACCCACTGGCGATGTCCCGTGCAGCCACCGGCTGTGTGGCCTGT 1260
Db	3300 GGACTCTTCAACACCCACTGGCGATGTCCCGTGCAGCCACCGGCTGTGTGGCCTGT 3359
QY	1261 GGTGCTGTGGCAGGCATGGGCGGGCCAGGGAGAAAGCAGGCTTTCAGGAGCAGTCCGCG 1320
Db	3360 GGTGCTGTGGCAGGCATGGGCGGGCCAGGGAGAAAGCAGGCTTTCAGGAGCAGTCCGCG 3419
QY	1321 GAGAGTGACGAGGAGGCGGCGCACGCTGCCTGTTCCTGTATGTGACCCAGTTTGTG 1380
Db	3420 GAGAGTGACGAGGAGGCGGCGCACGCTGCCTGTTCCTGTATGTGACCCAGTTTGTG 3479
QY	1381 TCCAGCCAGGCTTGGCAGAGTGAGCACTGCAATGCACCAAGTCTGGGTCAAGTTTGAT 1440
Db	3480 TCCAGCCAGGCTTGGCAGAGTGAGCACTGCAATGCACCAAGTCTGGGTCAAGTTTGAT 3539
QY	1441 ATCCGGGGGCACTGCCCTTGCAGAGCTGATGCCCGGGATGGGCCCCCGGGGATGCAGGC 1500
Db	3540 ATCCGGGGGCACTGCCCTTGCAGAGCTGATGCCCGGGATGGGCCCCCGGGGATGCAGGC 3599
QY	1501 CAGCAGAAAGGAATCAACACAGAAAACGCCCCAACTCCACAACCTTCTGTCAATGGCGAC 1560
Db	3600 CAGCAGAAAGGAATCAACACAGAAAACGCCCCAACTCCACAACCTTCTGTCAATGGCGAC 3659
QY	1561 ACCACAGGACCAAGAGCATCAAAAGAGGAGACCCCGATTCGGTGAGACCCAGCAGAG 1620
Db	3660 ACCACAGGACCAAGAGCATCAAAAGAGGAGACCCCGATTCGGTGAGACCCAGCAGAG 3719
QY	1621 GACCGTGTGGCCGAGGCCCTGCTTGTCTTCTCTGTGGAATGCTGGCTTCTTACC 1680
Db	3720 GACCGTGTGGCCGAGGCCCTGCTTGTCTTCTCTGTGGAATGCTGGCTTCTTACC 3779
QY	1681 GCGGTCAAACCTCTGCTTGGGCCATGAGCGAATACACATGGCCCTTCGCCCCCTCACTCCG 1740
Db	3780 GCGGTCAAACCTCTGCTTGGGCCATGAGCGAATACACATGGCCCTTCGCCCCCTCACTCCG 3839
QY	1741 GCCCTGCCAGTGATGACCGCATCAACAACATCTTGGACAGCATTATCGCACAGGTGGTG 1800
Db	3840 GCCCTGCCAGTGATGACCGCATCAACAACATCTTGGACAGCATTATCGCACAGGTGGTG 3899
QY	1801 GAACGGAAGATCCAGGAAAGCCCTGGGGCCGGGCTTCGAGCTGGCCCGGCTGTGCGC 1860
Db	3900 GAACGGAAGATCCAGGAAAGCCCTGGGGCCGGGCTTCGAGCTGGCCCGGCTGTGCGC 3959
QY	1861 AAGGGCTGGGCTGCCCTCTCTCCAGTGGGCCCGGCTGCTTCCCCCAGGSGCTTG 1920
Db	3960 AAGGGCTGGGCTGCCCTCTCTCCAGTGGGCCCGGCTGCTTCCCCCAGGSGCTTG 4019
QY	1921 CTGTGGCTGCAGGAGCCCCAGCCTTGGCCCTCGGCGTGGCTTCCACCTTTCAGGAGCAC 1980
Db	4020 CTGTGGCTGCAGGAGCCCCAGCCTTGGCCCTCGGCGTGGCTTCCACCTTTCAGGAGCAC 4079

Db 1981 TGGAGGCGAGGCGCAGCTGTGTTGGTGTAGGGATCCAAAGGACATTGCAGGGCAACCTG 2040
QY 2041 TGGGGACAGAAAGCTCTTGGGGCACTTGGAGGCCAGGTGCAGGCGCTGAGCCCTCGGA 2100
Db 2041 TGGGGACAGAAAGCTCTTGGGGCACTTGGAGGCCAGGTGCAGGCGCTGAGCCCTCGGA 2100
QY 2101 CCTCCCCAGCCCCAGAGCTGGGCAGCACAATTCCTCTGGGAGGGCTTCTCTGGGCTGAG 2160
Db 2101 CCTCCCCAGCCCCAGAGCTGGGCAGCACAATTCCTCTGGGAGGGCTTCTCTGGGCTGAG 2160
QY 2161 CTTGCCCCAAAGTCCAGACAGGGCTCTGTCTCTCTGTCGACCGAGCTTTGGGGGATGAG 2220
Db 2161 CTTGCCCCAAAGTCCAGACAGGGCTCTGTCTCTCTGTCGACCGAGCTTTGGGGGATGAG 2220
QY 2221 GACACAGCAGGGTGGAGAACTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCCTC 2280
Db 2221 GACACAGCAGGGTGGAGAACTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGCCCTC 2280
QY 2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGCGCTTCCCTGGTCCACTGGAG 2340
Db 2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGCGCTTCCCTGGTCCACTGGAG 2340
QY 2341 CCCAGCTCTGGGCAGCCATAGTGTGAGCCCGCACCGGGGACACCTGGGGACCAAGAAC 2400
Db 2341 CCCAGCTCTGGGCAGCCATAGTGTGAGCCCGCACCGGGGACACCTGGGGACCAAGAAC 2400
QY 2401 CTCTGTGTGGAGGTGGCCGACCTGGTTCAGCATCTCTGTGTCATGCCGACACACACTGCCT 2460
Db 2401 CTCTGTGTGGAGGTGGCCGACCTGGTTCAGCATCTCTGTGTCATGCCGACACACACTGCCT 2460
QY 2461 GCCTGGCACCGGGCACAGAAAGACTTCTTTTTCAGGCTTGGACGGGGGCTCTGGTCT 2520
Db 2461 GCCTGGCACCGGGCACAGAAAGACTTCTTTTTCAGGCTTGGACGGGGGCTCTGGTCT 2520
QY 2521 CCGGGCAGCCAGGTTCAGCACTGTGTGGCACGTGTTCCGGGCACAGGACGCCAGCGCATC 2580
Db 2521 CCGGGCAGCCAGGTTCAGCACTGTGTGGCACGTGTTCCGGGCACAGGACGCCAGCGCATC 2580
QY 2581 CGCCGCTTCTCCAGATGGTGTGCCCCGGCGGGCAGGCGCCCTGGAGCCTGGCGCCCCA 2640
Db 2581 CGCCGCTTCTCCAGATGGTGTGCCCCGGCGGGCAGGCGCCCTGGAGCCTGGCGCCCCA 2640
QY 2641 GGAGCTGTACTCTGGATGCAGGGCTCGCGCGCGCGCTGCGGAGGAGTGGGGCGTGAGC 2700
Db 2641 GGAGCTGTACTCTGGATGCAGGGCTCGCGCGCGCGCTGCGGAGGAGTGGGGCGTGAGC 2700
QY 2701 TGCTGACCCCTGTCTCAGGCCCCCGGAGAGGCGCTGTGTGCTGCTGACGGGCTCCCCAC 2760
Db 2701 TGCTGACCCCTGTCTCAGGCCCCCGGAGAGGCGCTGTGTGCTGCTGACGGGCTCCCCAC 2760
QY 2761 CAGGTGAGGGCCTGGTGAGCACAGTCAGCGTCACTCAGCACTTCTCTCCCCTGAGACC 2820
Db 2761 CAGGTGAGGGCCTGGTGAGCACAGTCAGCGTCACTCAGCACTTCTCTCCCCTGAGACC 2820
QY 2821 TCTGCCCTCTCTGCTCAGCTCTGCCACAGGGACCCAGCCTTCCCTCTGACTGCCACCTG 2880
Db 2821 TCTGCCCTCTCTGCTCAGCTCTGCCACAGGGACCCAGCCTTCCCTCTGACTGCCACCTG 2880
QY 2881 CTTTATGCCCAGATGGACTGGGCTGTGTTTCCAAGCAGTGAAGTGGCGTGGGACATTA 2940
Db 2881 CTTTATGCCCAGATGGACTGGGCTGTGTTTCCAAGCAGTGAAGTGGCGTGGGACATTA 2940
QY 2941 CAGGAGCCAAATAGAGGATGCTAGGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG 3000
Db 2941 CAGGAGCCAAATAGAGGATGCTAGGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG 3000
QY 3001 GTGCTCAGGCCAGGCACAACCTTTCAGAGGGGATGGCGTAGGGACTTGGGGATTTCTGG 3060
Db 3001 GTGCTCAGGCCAGGCACAACCTTTCAGAGGGGATGGCGTAGGGACTTGGGGATTTCTGG 3060
QY 3061 TCAACCCCAAGCACCACTCTGGGCACAGCAGGGCAGTCTGTTCCTCTCCCCTTAAG 3120

Db 3061 TCAACCCCAAGCACCACTCTGGGCACAAAGCAGGGCAGTCTGTTCCTCTCCCCTTAAG 3120
QY 3121 CCAACAACCAAGTCCACCAAGCTCACACCTGTCTTCTCAGGCTGGCATCTCCCCAC 3180
Db 3121 CCAACAACCAAGTCCACCAAGCTCACACCTGTCTTCTCAGGCTGGCATCTCCCCAC 3180
QY 3181 CCTGTGCCCTTTTATGTACAGG 3202
Db 3181 CCTGTGCCCTTTTATGTACAGG 3202
RESULT 4
ABX13987
ID ABX13987 standard; cDNA; 5709 BP.
XX
AC ABX13987;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human hairless protein cDNA.
KW Human; gene; ss; catalytic DNA; catalytic RNA; hairless protein;
KW hair loss; atrichia; hair growth; hirsutism; catalytic nucleic acid;
KW ribozyme; DNazyme; self-catalytic; hammerhead ribozyme; deoxy-ribozyme;
KW catalytic core; cleavage site; pharmaceutical; hair production;
KW hair follicle; anagen phase; catagen phase; hair removal product;
KW depilatory.
XX
OS Homo sapiens.
XX
PH Key
FT CDS
FT 1485..5054
FT /*tag= a
FT /product= "Hairless protein"
FT /transl_except= (pos:2820..2822,aa:Leu)
FT /transl_except= (pos:3234..3236,aa:Ser)
XX
PN WO200283891-A2.
XX
PD 24-OCT-2002.
XX
PF 12-APR-2002; 2002WO-US011683.
XX
PR 13-APR-2001; 2001US-0283618P.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Christiano AM;
XX
DR WPI; 2003-093020/08.
DR P-PSDB; ABG72775.
XX
PT New catalytic nucleic acid molecule that specifically cleaves Hairless
PT Protein mRNA, useful for inhibiting hair production by a hair-producing
PT cell, or for inhibiting transition of a hair follicle from anagen phase
PT to catagen phase.
XX
PS Claim 12; Fig 1; 65pp; English.
XX
CC The invention discloses a new catalytic DNA or RNA molecule that
CC specifically cleaves, or inhibits expression of, Hairless Protein mRNA
CC which comprises a catalytic domain that cleaves mRNA at a defined
CC consensus sequence and binding domains contiguous with the 5' and 3' ends
CC of the catalytic domain. Lack of expression of the hairless gene due to
CC inherited mutations leads to the complete loss of hair, known as
CC atrichia. Abundant hair growth, hirsutism, can be improved by inhibiting
CC the genes promoting hair growth, and one way to get targeted, transient
CC gene suppression is through the use of catalytic nucleic acid technology,
CC including ribozymes and DNazymes. Ribozymes are RNA structures which have
CC a self-catalytic enzymatic function and sequence specific RNA binding
CC ability. Small DNA oligonucleotides that have a similar structure to the
CC hammerhead ribozyme, called deoxy-ribozymes or DNazymes, having a
CC catalytic core and two sequence specific arms. The deoxy-ribozymes have

QY 2161 CTTGCGCCAAAGTCAGACGAGGGCTCTGTCTCTCTGCTGCACCGAGCTTTGGGGATGAG 2220
Db |||||||
2161 CTTGCGCCAAAGTCAGACGAGGGCTCTGTCTCTCTGCTGCACCGAGCTTTGGGGATGAG 2220
QY 2221 GACACCAGCAGGGTGGAGAACTAGCTGCCAGTCTGCCAGTCTCCGGAGTACTGCGCCCTC 2280
Db |||||||
2221 GACACCAGCAGGGTGGAGAACTAGCTGCCAGTCTGCCAGTCTCCGGAGTACTGCGCCCTC 2280
QY 2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCTTGCCTCGCTGCCTGAG 2340
Db |||||||
2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCTTGCCTCGCTGCCTGAG 2340
QY 2341 CCCAGCTCTGGGCAGCCTATGTTGTGAGCCCGCACCGGGGACACCTGGGGACCAAGAAC 2400
Db |||||||
2341 CCCAGCTCTGGGCAGCCTATGTTGTGAGCCCGCACCGGGGACACCTGGGGACCAAGAAC 2400
QY 2401 CTCTGTGTGAGGTGGCCGACCTGGTCAGCATCTCTGTGTCATGCCGACACACCACTGCCT 2460
Db |||||||
2401 CTCTGTGTGAGGTGGCCGACCTGGTCAGCATCTCTGTGTCATGCCGACACACCACTGCCT 2460
QY 2461 GCCTGGCACCGGGCACAGAAAGACTTCTTTTCAAGGCTTGGACGGGGAGGGGCTCTGGTCT 2520
Db |||||||
2461 GCCTGGCACCGGGCACAGAAAGACTTCTTTTCAAGGCTTGGACGGGGAGGGGCTCTGGTCT 2520
QY 2521 CCGGGCAGCCAGGTTCAGCACTGTGTGGCACGTGTTCCGGGCACAGGACGCCAGCGCATC 2580
Db |||||||
2521 CCGGGCAGCCAGGTTCAGCACTGTGTGGCACGTGTTCCGGGCACAGGACGCCAGCGCATC 2580
QY 2581 CGCCGCTTTCTCCAGATGTTGCCCGGGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2640
Db |||||||
2581 CGCCGCTTTCTCCAGATGTTGCCCGGGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2640
QY 2641 GGCAGCTGCTACCTGGATGCAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2700
Db |||||||
2641 GGCAGCTGCTACCTGGATGCAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2700
QY 2701 TGCTGGACCCCTGCTCCAGGCCCCCGGAGAGGCCGCTGCTGGTGCCTGCAGGGGCTCCCCAC 2760
Db |||||||
2701 TGCTGGACCCCTGCTCCAGGCCCCCGGAGAGGCCGCTGCTGGTGCCTGCAGGGGCTCCCCAC 2760
QY 2761 CAGTGCAGGGCCTGTTGAGCACAGTCAGCGTCACTCAGCACTTCTCTCCCTGAGACC 2820
Db |||||||
2761 CAGTGCAGGGCCTGTTGAGCACAGTCAGCGTCACTCAGCACTTCTCTCCCTGAGACC 2820
QY 2821 TCTGCCCTCTCTGCTCAGCTCTGCCACAGGGACCCAGCCCTTCCCTTACTGCCACCTG 2880
Db |||||||
2821 TCTGCCCTCTCTGCTCAGCTCTGCCACAGGGACCCAGCCCTTCCCTTACTGCCACCTG 2880
QY 2881 CTTTATGCCCAGATGGACTGGCTGTGTTCAGCAGTGAAGTGGCCGTGGGGACATTA 2940
Db |||||||
2881 CTTTATGCCCAGATGGACTGGCTGTGTTCAGCAGTGAAGTGGCCGTGGGGACATTA 2940
QY 2941 CAGGAGSCAAATAGAGGGATGCTAGGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG 3000
Db |||||||
2941 CAGCAGSCAAATAGAGGGATGCTAGGTGTCTGGGATCGGGGTGGGACAGGTAGACCAG 3000
QY 3001 GTGCTCAGCCAGGCACAACCTTCAGCAGGGGATGGCGCTAGGGGACTTGGGATTCTGG 3060
Db |||||||
3001 GTGCTCAGCCAGGCACAACCTTCAGCAGGGGATGGCGCTAGGGGACTTGGGATTCTGG 3060
QY 3061 TCAACCCACAAAGCACCACCTCTGGGCAAGCAGGGGCACTCTGTTCCCTTCCCTTAA 3120
Db |||||||
3061 TCAACCCACAAAGCACCACCTCTGGGCAAGCAGGGGCACTCTGTTCCCTTCCCTTAA 3120
QY 3121 CCAACACACAGTGCACCAAGCTCACCTGTCTCTCAGGTGGCATCTCCCCAC 3180
Db |||||||
3121 CCAACACACAGTGCACCAAGCTCACCTGTCTCTCAGGTGGCATCTCCCCAC 3180
QY 3181 CCTGTGCCCTTTTATGTACAGG 3202
Db |||||||
3181 CCTGTGCCCTTTTATGTACAGG 3202

RESULT 3

ACD26351
ID ACD26351 standard; cDNA; 3202 BP.

XX
AC ACD26351;

XX
DT 08-SEP-2003 (first entry)

XX
DE Human Hairless (HR) cDNA.

XX
KW Human; hairless; HR; gene; ss; thyroid hormone receptor; hair growth;
hair loss; family pedigree; radiation hybrid; somatic cell hybrid.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT CDS 1. .2955

FT /*tag= a

FT /product= "Human hairless protein"

FT /partial

FT /transl_except= (pos:148. .150, aa:Met)

FT /transl_except= (pos:418. .420, aa:Glu)

FT /transl_except= (pos:1150. .1152, aa:Val)

FT /transl_except= (pos:1336. .1338, aa:Gln)

FT /transl_except= (pos:1399. .1401, aa:Gln)

FT /transl_except= (pos:1585. .1590, aa:Gln-Gln)

FT /transl_except= (pos:1807. .1809, aa:Lys)

FT /transl_except= (pos:1816. .1818, aa:Gln)

FT /transl_except= (pos:1933. .1935, aa:Gln)

FT /transl_except= (pos:1975. .1977, aa:Gln)

FT /transl_except= (pos:2239. .2241, aa:Asn)

FT /note= "No start codon shown"

XX US2003027300-A1.

XX
PN 06-FEB-2003.

XX
PD 21-DEC-2001; 2001US-00024368.

XX
PF 07-APR-1998; 98US-0080888P.

XX
PR 07-APR-1999; 99US-00287354.

XX
PA (CARN-) CARNEGIE INST WASHINGTON.

XX
PI Thompson CC;

XX
PS WPI; 2003-492034/58.

XX
DR P-PSDB; ABU62540.

XX
CC New Hairless polypeptide and polynucleotide, useful for identifying and
detecting this genetic marker in family pedigrees or human-rodent somatic
cell hybrids, or detecting interacting proteins that bind hairless gene
or protein.

XX
PS Claim 3; Page 14-17; 34pp; English.

XX
CC The invention relates to a human hairless (HR) polynucleotide encoding
the hairless (Hr) polypeptide. The invention also relates to an
expression system comprising an expression construct which produces a
polypeptide with hairless transcription factor activity, a reporter
construct comprising a transcription regulatory region responsive to
hairless transcription activity to regulate transcription of the reporter
gene which is mediated by the transcription regulatory region and methods
of screening for chemical agents which modulate hairless-mediated
transcription, binding between hairless and thyroid hormone receptor or
hairless activity. The polynucleotide is useful as a probe or primer to
quantitate cognate RNA and DNA within cells, which can be subsequently
used to correlate hair growth or loss with hairless expression or
hairless-regulated transcription. The amino acid sequence of hairless
antigen can be used for preparing specific binding molecules (e.g.
polyclonal or monoclonal antibodies) for monitoring protein expression,
for affinity purification and for functional studies. The human hairless

Best Local Similarity 99.8%; Pred. No. 0;									
Matches 3196; Conservative 0; Mismatches 6; Indels 0; Gaps 0;									
QY	1	TTTTACTAC	AAGATCC	GAGCAT	TCCAG	TTGGCAA	AGGAGCC	CTTTGGC	AGCTGCGGAA 60
Db	1	TTTTACTAC	AAGATCC	GAGCAT	TCCAG	TTGGCAA	AGGAGCC	CTTTGGC	AGCTGCGGAA 60
QY	61	CTTGGGTTG	TTGGCTTAA	ACTCTG	TGGCAC	CTTG	CAGAGAG	CCCGGGAG	CCCGAACGC 120
Db	61	CTTGGGTTG	TTGGCTTAA	ACTCTG	TGGCAC	CTTG	CAGAGAG	CCCGGGAG	CCCGAACGC 120
QY	121	CCTTCACTG	CACAGAGG	ATGGAG	ATGGAG	ATGGAG	ATGGAG	ATGGAG	ATGGAG 180
Db	121	CCTTCACTG	CACAGAGG	ATGGAG	ATGGAG	ATGGAG	ATGGAG	ATGGAG	ATGGAG 180
QY	181	CTCTTCTCT	GGGAGCC	AGACACTG	TGCTGG	CCCTGG	ACCTCT	CTGGCC	CCCTGG 240
Db	181	CTCTTCTCT	GGGAGCC	AGACACTG	TGCTGG	CCCTGG	ACCTCT	CTGGCC	CCCTGG 240
QY	241	CTTGTTCAT	ACTCTTGG	CAACG	CTCTGG	GGCTGG	AGCTGG	CCGCTGG	CAATCCTT 300
Db	241	CTTGTTCAT	ACTCTTGG	CAACG	CTCTGG	GGCTGG	AGCTGG	CCGCTGG	CAATCCTT 300
QY	301	CTGGGGCAC	AGCAAC	CAAGGTG	CCCTCT	CTCTG	AGCCGCT	GTCA	CCCGGGGC 360
Db	301	CTGGGGCAC	AGCAAC	CAAGGTG	CCCTCT	CTCTG	AGCCGCT	GTCA	CCCGGGGC 360
QY	361	TGCTGTTCA	TCTACCC	ACCTAA	AGGTGG	GATCTT	TGGCCCT	TGTGG	AAAGTGCCAG 420
Db	361	TGCTGTTCA	TCTACCC	ACCTAA	AGGTGG	GATCTT	TGGCCCT	TGTGG	AAAGTGCCAG 420
QY	421	GAGGGCTTG	AGGGGGT	GCCAGT	TGGAGCC	AGCAAC	CCAGCC	AGCAAGT	GAACAGGCC 480
Db	421	GAGGGCTTG	AGGGGGT	GCCAGT	TGGAGCC	AGCAAC	CCAGCC	AGCAAGT	GAACAGGCC 480
QY	481	TCTGGCCCC	AGGGCCTG	TCCCCC	AGCCAC	CAAGCTG	AAAGAC	ATGGCT	CAACA 540
Db	481	TCTGGCCCC	AGGGCCTG	TCCCCC	AGCCAC	CAAGCTG	AAAGAC	ATGGCT	CAACA 540
QY	541	CGGCACTCG	GAGCAGT	TTGAATG	TCCACG	GGTG	CCCTG	AGGTCG	AGGAGCCGGTT 600
Db	541	CGGCACTCG	GAGCAGT	TTGAATG	TCCACG	GGTG	CCCTG	AGGTCG	AGGAGCCGGTT 600
QY	601	GTCGGCTCC	GGGCCCTC	AAAAGGG	CAGGCAG	CCCCAG	AGTCC	AGGGAG	CAATGSSCAGT 660
Db	601	GTCGGCTCC	GGGCCCTC	AAAAGGG	CAGGCAG	CCCCAG	AGTCC	AGGGAG	CAATGSSCAGT 660
QY	661	CCAGCCCCC	AAAGCGGC	ACCTTTCC	AGGCAC	TGCAAG	AACTG	CAAG	AGGGGGT 720
Db	661	CCAGCCCCC	AAAGCGGC	ACCTTTCC	AGGCAC	TGCAAG	AACTG	CAAG	AGGGGGT 720
QY	721	TGSCAGGAG	GGTTG	CGGGAC	ACATCG	ATAGG	GAACA	AGGATG	TGGACAGCATGAT 780
Db	721	TGSCAGGAG	GGTTG	CGGGAC	ACATCG	ATAGG	GAACA	AGGATG	TGGACAGCATGAT 780
QY	781	GACGAGAA	AGGACCC	CAAGAT	TGGCC	AGGCC	AGTCTC	AGGACCTT	CAGGACATA 840
Db	781	GACGAGAA	AGGACCC	CAAGAT	TGGCC	AGGCC	AGTCTC	AGGACCTT	CAGGACATA 840
QY	841	CCATGCCCT	GCTTCTCC	CTGCAAA	ACTGG	CTCAAT	GCCAAA	AGTTGT	GCCCAGGCAGCTGGA 900
Db	841	CCATGCCCT	GCTTCTCC	CTGCAAA	ACTGG	CTCAAT	GCCAAA	AGTTGT	GCCCAGGCAGCTGGA 900
QY	901	GAGGAGGAG	GGGACG	CGCTCT	CAGCA	AGTGG	GAGATG	CGCTCT	TGGGAGGGGAG 960
Db	901	GAGGAGGAG	GGGACG	CGCTCT	CAGCA	AGTGG	GAGATG	CGCTCT	TGGGAGGGGAG 960
QY	961	CTGCAGCAG	GAGGAAG	ACACAG	CCACCA	CTCC	AGCTCT	GAGGA	AGGGTCCGGC 1020
Db	961	CTGCAGCAG	GAGGAAG	ACACAG	CCACCA	CTCC	AGCTCT	GAGGA	AGGGTCCGGC 1020
QY	1021	CCTGACAG	CCGGCTC	AGCACAG	GGCCTC	GGCAAG	CACTG	CTCAG	TGTTTGGGGACCGA 1080
Db	1021	CCTGACAG	CCGGCTC	AGCACAG	GGCCTC	GGCAAG	CACTG	CTCAG	TGTTTGGGGACCGA 1080

Db	1021	CCTGACAG	CCGGCTC	AGCACAG	GGCCTC	GGCAAG	CACTG	CTCAG	TGTTTGGGGACCGA 1080
QY	1081	CTGTGCCG	CCTGCT	CGGAGG	AGCGG	AGCCCT	TGGCTT	GGGCCC	AGCGGAAGGCCAA 1140
Db	1081	CTGTGCCG	CCTGCT	CGGAGG	AGCGG	AGCCCT	TGGCTT	GGGCCC	AGCGGAAGGCCAA 1140
QY	1141	GGGCCAG	CCGTGAC	AGGGGAC	AGCC	CCAGGCAT	TTCC	ACGCTG	CTGCAGCCGTTGCCACCAT 1200
Db	1141	GGGCCAG	CCGTGAC	AGGGGAC	AGCC	CCAGGCAT	TTCC	ACGCTG	CTGCAGCCGTTGCCACCAT 1200
QY	1201	GGACTCTT	CAACAC	CCCACT	TGGCGAT	GTCCCC	GTGC	AGCC	ACCGGCTGTGTGGCCTGT 1260
Db	1201	GGACTCTT	CAACAC	CCCACT	TGGCGAT	GTCCCC	GTGC	AGCC	ACCGGCTGTGTGGCCTGT 1260
QY	1261	GGTCTGTG	GCAGGC	ACTGGG	CGGCC	AGGAGAA	AGCAGG	CTTT	TCAGGAGCAGTCCGCG 1320
Db	1261	GGTCTGTG	GCAGGC	ACTGGG	CGGCC	AGGAGAA	AGCAGG	CTTT	TCAGGAGCAGTCCGCG 1320
QY	1321	GAGGAGTG	CACGC	AGGAGG	CCGG	CACGCTG	CTCT	TCTCTG	ATGCTGACCCAGTTTGTG 1380
Db	1321	GAGGAGTG	CACGC	AGGAGG	CCGG	CACGCTG	CTCT	TCTCTG	ATGCTGACCCAGTTTGTG 1380
QY	1381	TCCAGCC	AGGCTTT	TGGCAG	AGCTG	AGCAAT	GCAC	CCAGG	CTGGGTCAAGTTTGAT 1440
Db	1381	TCCAGCC	AGGCTTT	TGGCAG	AGCTG	AGCAAT	GCAC	CCAGG	CTGGGTCAAGTTTGAT 1440
QY	1441	ATCCGGGG	GCAC	TGCCCC	CTGCC	AAAGCTG	ATG	CCCC	GGGTATGGGCCCCCGGGGATGCAGGC 1500
Db	1441	ATCCGGGG	GCAC	TGCCCC	CTGCC	AAAGCTG	ATG	CCCC	GGGTATGGGCCCCCGGGGATGCAGGC 1500
QY	1501	CAGCAGA	AGGAAT	CAAC	CACAG	AAAA	CGCCCC	AACTCC	ACAACTTCTGCAATGGCGAC 1560
Db	1501	CAGCAGA	AGGAAT	CAAC	CACAG	AAAA	CGCCCC	AACTCC	ACAACTTCTGCAATGGCGAC 1560
QY	1561	ACCCACAG	GACCA	AGAGCA	TCAA	AGAGG	AGAC	CCCC	CGAATTCGGT
Db	1561	ACCCACAG	GACCA	AGAGCA	TCAA	AGAGG	AGAC	CCCC	CGAATTCGGT
QY	1621	GACCGTGT	GGCCG	AGGCG	CCCTG	CTCTT	CTCTG	CGAA	CTGCTGGCTTCTACC 1680
Db	1621	GACCGTGT	GGCCG	AGGCG	CCCTG	CTCTT	CTCTG	CGAA	CTGCTGGCTTCTACC 1680
QY	1681	GCGTCAAA	ACTCTG	CTTGG	SCATG	AGCGA	ATA	CACAT	TGCCCCCTCGCTCACTCCG 1740
Db	1681	GCGTCAAA	ACTCTG	CTTGG	SCATG	AGCGA	ATA	CACAT	TGCCCCCTCGCTCACTCCG 1740
QY	1741	GCCCTGCCC	AGTGAT	GAC	CCGAT	CACCA	ACAT	CCCTG	GACAGCAT
Db	1741	GCCCTGCCC	AGTGAT	GAC	CCGAT	CACCA	ACAT	CCCTG	GACAGCAT
QY	1801	GAACGGA	AGATCC	AGG	AGAA	AGCCCT	TGGG	CCCGGG	CTTCGAGCTGGCCCGGCTTCGCGC 1860
Db	1801	GAACGGA	AGATCC	AGG	AGAA	AGCCCT	TGGG	CCCGGG	CTTCGAGCTGGCCCGGCTTCGCGC 1860
QY	1861	AAGGGCTG	GGCCTG	CCCTCT	CTCC	AGTCC	GGCCCG	CGCTGC	CTCCCCCAGGGGCTTTG 1920
Db	1861	AAGGGCTG	GGCCTG	CCCTCT	CTCC	AGTCC	GGCCCG	CGCTGC	CTCCCCCAGGGGCTTTG 1920
QY	1921	CTGTGGCT	GACGAG	CCCG	CAGC	CTTG	CCCTCG	SGCTGG	CTTCCACGAGCAC 1980
Db	1921	CTGTGGCT	GACGAG	CCCG	CAGC	CTTG	CCCTCG	SGCTGG	CTTCCACGAGCAC 1980
QY	1981	TGGAGGC	AGGGCC	AGCCT	TGTG	TGTC	AGGAT	TCAA	AGGACATTG
Db	1981	TGGAGGC	AGGGCC	AGCCT	TGTG	TGTC	AGGAT	TCAA	AGGACATTG
QY	2041	TGGGGGAC	AGAGCT	TTGGG	GCACTT	GAGG	CCAGG	TCAG	CGCTGAGCCCGCTCGGA 2100
Db	2041	TGGGGGAC	AGAGCT	TTGGG	GCACTT	GAGG	CCAGG	TCAG	CGCTGAGCCCGCTCGGA 2100
QY	2101	CCTCCCC	AGCC	AGCAG	CCCTG	GGCAG	CAACA	ATT	TGGAGGGCTTCTCTGGCCTGAG 2160
Db	2101	CCTCCCC	AGCC	AGCAG	CCCTG	GGCAG	CAACA	ATT	TGGAGGGCTTCTCTGGCCTGAG 2160

Db 1921 CTGTGGCTGCAGAGCCCCAGCCTTGCCCTGGCGTGGCTTCCACCTCTTCCAGGAGCAC 1980

QY 1981 TGGAGGCAGGGCCAGCCTGTGTGGTGTAGGGATCCAAAGGACATTGCGAGGCAACCTTG 2040

Db 1981 TGGAGGCAGGGCCAGCCTGTGTGGTGTAGGGATCCAAAGGACATTGCGAGGCAACCTTG 2040

QY 2041 TGGGGGACAGAAAGCTCTTGGGGCACTTGGAGGCCAGGTGCAGGCGCTGAGCCCTCGGA 2100

Db 2041 TGGGGGACAGAAAGCTCTTGGGGCACTTGGAGGCCAGGTGCAGGCGCTGAGCCCTCGGA 2100

QY 2101 CCTCCCCAGCCCAGCAGCTTGGGCAGCACAATCTGCGAGGGCTTCTCCTGGCCTGAG 2160

Db 2101 CCTCCCCAGCCCAGCAGCTTGGGCAGCACAATCTGCGAGGGCTTCTCCTGGCCTGAG 2160

QY 2161 CTTGCCCCAAAGTCAGACGAGGGCTCTGTCTCCTGCTGCACCGAGCTTTGGGGGATGAG 2220

Db 2161 CTTGCCCCAAAGTCAGACGAGGGCTCTGTCTCCTGCTGCACCGAGCTTTGGGGGATGAG 2220

QY 2221 GACACAGCAGGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGGCCCTC 2280

Db 2221 GACACAGCAGGGTGGAGAACCTAGCTGCCAGTCTGCCACTTCCGGAGTACTGCGGCCCTC 2280

QY 2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCTTCCCTGCGTCCACTGGAG 2340

Db 2281 CATGGAATACTCAACCTGGCTTCTTACCTCCACCGGGCTTCCCTGCGTCCACTGGAG 2340

QY 2341 CCCAGCTCTGGGCAGCCTATGGTGTGAGCCCGCACCCGGGACACCTGGGGACCAAGAAC 2400

Db 2341 CCCAGCTCTGGGCAGCCTATGGTGTGAGCCCGCACCCGGGACACCTGGGGACCAAGAAC 2400

QY 2401 CTCTGTGTGGAGTGGCCGACCTGGTCAGCATCTCTGGTGCATGCCGACACACCACTGCGT 2460

Db 2401 CTCTGTGTGGAGTGGCCGACCTGGTCAGCATCTCTGGTGCATGCCGACACCACTGCGT 2460

QY 2461 GCCTGGCACCGGGCACAGAAAGACTTCCTTTTCAGGCCCTGGACGGGAGGGGCTCTGGTCT 2520

Db 2461 GCCTGGCACCGGGCACAGAAAGACTTCCTTTTCAGGCCCTGGACGGGAGGGGCTCTGGTCT 2520

QY 2521 CCGGCAGCCAGGTGAGCATCTGTGGCACGTGTTCCGGGCACAGGACGCCCCAGCGCATC 2580

Db 2521 CCGGCAGCCAGGTGAGCATCTGTGGCACGTGTTCCGGGCACAGGACGCCCCAGCGCATC 2580

QY 2581 CGCGCTTTCTCCAGATGGTGTGCCCGCCCGGGCAGGCGCCTGGAGCCTGGCGCCCCA 2640

Db 2581 CGCGCTTTCTCCAGATGGTGTGCCCGCCCGGGCAGGCGCCTGGAGCCTGGCGCCCCA 2640

QY 2641 GGCAGTGTCTACCTGGATGCAGGGCTGCGGCGCGCCCTGCGGGAGGAGTGGGCGTGAGC 2700

Db 2641 GGCAGTGTCTACCTGGATGCAGGGCTGCGGCGCGCCCTGCGGGAGGAGTGGGCGTGAGC 2700

QY 2701 TGTGACCCCTGCTCCAGGCCCCCGAGAGGCGCGTGTGCTGCTGCAGGGGCTCCCCAC 2760

Db 2701 TGTGACCCCTGCTCCAGGCCCCCGAGAGGCGCGTGTGCTGCTGCAGGGGCTCCCCAC 2760

QY 2761 CAGGTGACGGCCCTGGTGAGCACAGTCAGCGTCACTCAGCATCTTCTCTCCCTGAGACC 2820

Db 2761 CAGGTGACGGCCCTGGTGAGCACAGTCAGCGTCACTCAGCATCTTCTCTCCCTGAGACC 2820

QY 2821 TCTGCCCTCTCTGCTCAGCTCTGCCACAGGACCCAGCCTTCCCTGACTGCCACCTG 2880

Db 2821 TCTGCCCTCTCTGCTCAGCTCTGCCACAGGACCCAGCCTTCCCTGACTGCCACCTG 2880

QY 2881 CTTTATGCCAGATGGACTGGGCTGTGTTCGAAGCAGTGAAGTGGCGTGGGACATTA 2940

Db 2881 CTTTATGCCAGATGGACTGGGCTGTGTTCGAAGCAGTGAAGTGGCGTGGGACATTA 2940

QY 2941 CAGGAGGCCAAATAGAGGGATGTAGGTGTCTGGGATCGGGTGGGACAGGTAGACCAG 3000

Db 2941 CAGGAGGCCAAATAGAGGGATGTAGGTGTCTGGGATCGGGTGGGACAGGTAGACCAG 3000

QY 3001 GTGCTCAGCCCAGGCACAACCTTCAGCAGGGGATGGCGCTAGGGGACTTGGGATTTCTGG 3060

Db 3001 GTGCTCAGCCCAGGCACAACCTTCAGCAGGGGATGGCGCTAGGGGACTTGGGATTTCTGG 3060

QY 3061 TCAACCCCCACAAGCACCACTCTGGGCACAAGCAGGGCACTCTGTTCCCTCCCTTTAAG 3120

Db 3061 TCAACCCCCACAAGCACCACTCTGGGCACAAGCAGGGCACTCTGTTCCCTCCCTTTAAG 3120

QY 3121 CCAACAACCCACAGTGCCACCAAGCTCACACCTGTCTCTCAGGCTGGCATCTCCCCAC 3180

Db 3121 CCAACAACCCACAGTGCCACCAAGCTCACACCTGTCTCTCAGGCTGGCATCTCCCCAC 3180

QY 3181 CCTGTGCCCTTTTATGTACAGG 3202

Db 3181 CCTGTGCCCTTTTATGTACAGG 3202

RESULT 2

AAD31077

ID AAD31077 standard; DNA; 3202 BP.

XX

AC AAD31077;

XX

DT 18-JUN-2002 (first entry)

XX

DE Human Hairless gene (HR).

XX

KW Human; Hairless protein; Hr protein; HR gene; dermatological condition; hair loss; gene therapy; ds.

XX

OS Homo sapiens.

XX

PH Key

FT CDS

FT 1. .2955

FT /*tag= a

FT /product= "Human Hr protein"

FT /transl_except= (pos:2932. .2934, aa:Gly)

FT /transl_except= (pos:2944. .2946, aa:Glu)

FT /note= "CDS does not include start codon"

FT /partial

XX

PN US6348348-B1.

XX

PD 19-FEB-2002.

XX

PF 07-APR-1999; 99US-00287354.

XX

PR 07-APR-1998; 98US-0080888P.

XX

PA (CARN-) CARNEGIE INST WASHINGTON.

XX

PI Thompson CC;

XX

DR WPI; 2002-204622/26.

DR P-PSDB; AAE19794.

XX

PT Novel expression construct, useful in the diagnosis and treatment of dermatological conditions, such as hair loss, contains a Hairless gene sequence.

XX

PS Claim 4; Col 25-32; 48pp; English.

XX

CC The invention relates to human Hairless (Hr) polypeptides and nucleic acid molecules (HR) encoding such polypeptides. The invention also relates to the Hairless expression constructs which may be used in transcription assays. Sequences of the invention are used in diagnosis and treatment of dermatological conditions such as hair loss. They are also used in gene therapy. Polynucleotides of the invention can be used as probes for the detection of hair loss. The present sequence is human Hairless gene (HR). Note: This sequence is stated as being the same as that shown as SEQ ID NO:1 in Column 35-38 (AAD31072) of the specification. However the sequences differ at positions

XX

SQ Sequence 3202 BP; 599 A; 1044 C; 1026 G; 533 T; 0 U; 0 Other;

Query Match 99.7%; Score 3192.4; DB 6; Length 3202;